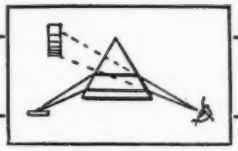
Journal of the OPTICAL SOCIETY *of* AMERICA

VOLUME 24
No. 3



MARCH
1934

CONTENTS

Current Advances in Photographic Photometry . . GEORGE R. HARRISON 59

Development of the Photographic Objective . . . . . R. KINGSLAKE 73

Effect, on the Latent Image, of Post-Heating and of Melting the Gelatin. . . . . . A. J. REARDON AND H. P. GRIGGS, S.J. 85

Some Comments on the Ives Theory of Flicker . . . . PERCY W. COBB 91

Published monthly
at Prince and Lemon Sts., Lancaster, Pa.
by the
AMERICAN INSTITUTE OF PHYSICS
Incorporated
11 East 38th Street, New York, N. Y.
FOR AND UNDER THE EDITORIAL SUPERVISION OF
THE OPTICAL SOCIETY OF AMERICA

THE OPTICAL SOCIETY OF AMERICA

OFFICERS OF THE SOCIETY

W. B. RAYTON, *President*
Bausch & Lomb Optical Company
Rochester, New York

ARTHUR C. HARDY, *Vice-President*
Massachusetts Institute of Technology
Cambridge, Massachusetts

L. B. TUCKERMAN, *Secretary*
Bureau of Standards
Washington, D. C.

HENRY F. KURTZ, *Treasurer*
Bausch & Lomb Optical Company
Rochester, New York

The Board of Directors consists of the above officers; the Editor of the Journal; the Past-President; and five elected members:

E. C. CRITTENDEN, *Past-President*
Bureau of Standards
Washington, D. C.

F. K. RICHTMYER, *Editor*
Cornell University
Ithaca, New York

Elected Members

ARTHUR H. COMPTON
Ryerson Physical Laboratory
University of Chicago
Chicago, Illinois

HERBERT E. IVES
463 West Street
New York, New York

CARL W. KEUFFEL
Keuffel & Esser Company
Hoboken, New Jersey

GEORGE B. PEGRAM
Columbia University
New York, New York

R. C. GIBBS
Department of Physics
Cornell University
Ithaca, New York

THE OPTICAL SOCIETY OF AMERICA was organized in 1916 "*to increase and diffuse the knowledge of optics in all its branches, pure and applied, to promote the mutual interests of investigators of optical problems, of designers, manufacturers and users of optical instruments and apparatus of all kinds, and to encourage cooperation among them.*" It aims to serve the interests of all who are engaged in any branch of optics from fundamental research to the manufacture of optical goods.

MEMBERSHIP

Regular Members: "Any person who has contributed materially to the advancement of optics is eligible for regular membership."

Associate Members: "Any person who is interested in optics is eligible for associate membership."

Corporate Members: "Any corporation desirous of encouraging the work of the Society is eligible for corporation membership."

The Society cordially invites to membership all who are interested in any branch of optics, either in research, in instruction, in optical or illuminating engineering, in the manufacture and distribution of optical goods of all kinds, or in physiological and medical optics.

ANNUAL DUES

Regular members, $7.50; Associate members, $5.00; Corporate members, $50.00. These dues include subscription to the Society's official journal, the Journal of the Optical Society of America.

MEETINGS

The Society holds two meetings per year: An annual meeting in the autumn, and a midwinter meeting jointly with the American Physical Society.

For further information concerning the Optical Society of America address the Secretary, Dr. L. B. Tuckerman, Bureau of Standards, Washington, D. C.

ROCHESTER, NEW YORK SECTION OF THE OPTICAL SOCIETY OF AMERICA

C. E. BAHN, *President*
MRS. H. G. KINGSLAKE, *Secretary*

WALTER CLARK, *Vice-President*
A. A. LEVINE, *Treasurer*

Councilors: GUSTAVE FASSIN, OTTO SANDVIK, C. J. STAUD, T. R. WILKINS

Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 24 MARCH, 1934 NUMBER 3

Current Advances in Photographic Photometry*

GEORGE R. HARRISON, *George Eastman Research Laboratory of Physics, Massachusetts Institute of Technology*
(Received December 16, 1933)

TABLE OF CONTENTS

INTRODUCTION 59
§1. METHODS OF PLATE CALIBRATION 60
TABLE OF METHODS USED IN VARYING INTENSITIES FOR CALIBRATING PLATES 60
§2. DISCUSSION OF CALIBRATION METHODS OF TABLE I 62
§3. SPECIAL EYE-MATCH HOMOCHROMATIC METHODS FOR ABSORPTION SPECTROPHOTOMETRY 66
§4. HETEROCHROMATIC PHOTOMETRY 68
§5. THE EXTREME ULTRAVIOLET 69
§6. DENSITOMETERS AND MICROPHOTOMETERS 69
BIBLIOGRAPHY 71

IN an article published four years ago[1] the writer discussed the instruments and methods then used for measuring spectral light intensities by photography. While few fundamental additions to the art of photographic photometry have been made since that paper was written, a number of articles have been published which settle points then uncertain and suggest further advances, and it seems desirable to present a more up-to-date summary. The present paper is merely an extension of the previous one, which takes up a number of points not discussed in this.

The situation regarding choice of spectrograph, emulsion, method of timing, and process of development remains sensibly as in 1929. The chief advances have been in methods of calibrating the emulsion response to variations of intensity, and in the design of microphotometers for measuring density.

We divide the discussion as before into sections dealing with homochromatic photometry, or comparison of beams of light of the same wave-length but of differing intensity, and heterochromatic photometry, in which beams of differing wave-length and intensity are to be compared. In the homochromatic category we again distinguish between cases involving close-lying groups of lines such as Zeeman, Stark, or hyperfine structure patterns, and those in which each wave-length is to be compared with itself reduced in intensity, as by absorption, reflection, or scattering.*

* Invited paper presented at the symposium on *One Hundred Years of Photography* at the Buck Hill Falls meeting of the Optical Society of America, October, 1933.

* A comprehensive survey of the methods of photographic photometry used by the Utrecht workers is given in the book *Objektive Spektralphotometrie* by Ornstein, Moll and Burger (Vieweg, 1932).

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

TABLE I. *Methods of imposing intensity calibration marks on photometric plates.*

		Selective Neutral	Requirements for use	Remarks
Single Exposure Methods	Auto-calibration	Step-sector disk / Logarithmic sector	At slit, rapid rotation, stigmatic spectrograph.	Convenient and accurate if made and used accurately.
		Step-weakener / Wedge-weakener	At slit, calibrate, stigmatic spectrograph.	Convenient and accurate once calibrated at proper wave-lengths.
		Step-diaphragm / Wedge-diaphragm	Astigmatic optical system, careful adjustment.	Convenient but unsuitable for accurate work if neutrality is assumed.
		Step-diaphragm / Wedge-diaphragm	Rowland type grating mounting.	Convenient for large gratings, but covers only short range at a setting.
		Step-weakener / Wedge-weakener	At plate, astigmatic spectrograph or long even lines.	Very accurate and simple, once suitable weakeners obtained and calibrated.
		Theoretical astigmatism	Paschen or other fixed grating mount.	Not yet thoroughly tested.
		Theoretical line shapes	Large instrumental broadening.	Seldom used except for interpreting old spectrograms without intensity marks.
		Coarse crossed grating	Calibrate grating orders for intensity.	Useful in astronomical work or qualitative estimates.
		Geometrical shadow	Special cases only.	Seldom used.
	Auxiliary Source	Step-slit / Wedge-slit	Continuous spectrum, stigmatic spectrograph.	Easy but questionable unless Eberhard effect avoided.
		Theoretical line ratios	Avoid self-reversal and excitation errors.	Useful when all else fails.
Multiple Exposure Methods	Steady Auxiliary Source	Variable diaphragms	Uniform beam section.	Careful tests for uniformity needed, usually less convenient than others.
		Wire gauze screens	Wide optical beam.	Avoid diffraction errors.
		Semi-neutral filters	Calibrate filters for transmission *vs.* wave-length.	Avoid refraction errors.
		Inverse square law	Long optical path.	Suitable scatterer needed, avoid atmospheric absorption.
		Variable slit length	Astigmatic spectrograph.	Simple if illumination carefully controlled.
		Variable slit width	Continuous spectrum.	Avoid diffraction errors and Eberhard effect.
		Diffuse reflection	Calibrate reflector.	Seldom used.
		Variable exposure time	Compensate reciprocity failure.	Often easiest method for qualitative work; to be used with discrimination.
		Variable sector disk	At slit, rapid rotation.	Avoid errors from extreme conditions.
		Control of source	Suitable calibrated source.	Seldom convenient.
		Polarizers	Transparent polarizers.	Seldom convenient for photography.

§1. METHODS OF PLATE CALIBRATION

When faced with a problem involving photographic photometry, one must first decide which method of imposing intensity calibration marks on the plate will be most suitable. There is no general answer to this question, since a method applicable to all cases has not yet been found. We shall consider the various methods which have been proposed and discuss the merits and disadvantages of each; by consulting such a list as is given in Table I, one should be able to select the most suitable method available for any given case.

Ideally, the spectrum to be studied should be used as its own calibrating spectrum, and this can be done in some cases where we can control the illumination along each line by having various sections of the line reduced in intensity by known amounts. This is not only a great time saver, since only one exposure is needed for calibration, but automatically insures constancy of exposure time. Also, the Eberhard effect is reduced insofar as calibration is concerned, and a number of determinations of each intensity ratio desired can be obtained from each pair of lines. When two lines being compared have the same wave-length their two calibration curves must be similar, but are displaced due to the different intensities of the lines, and the amount of this displacement gives the intensity ratio desired.

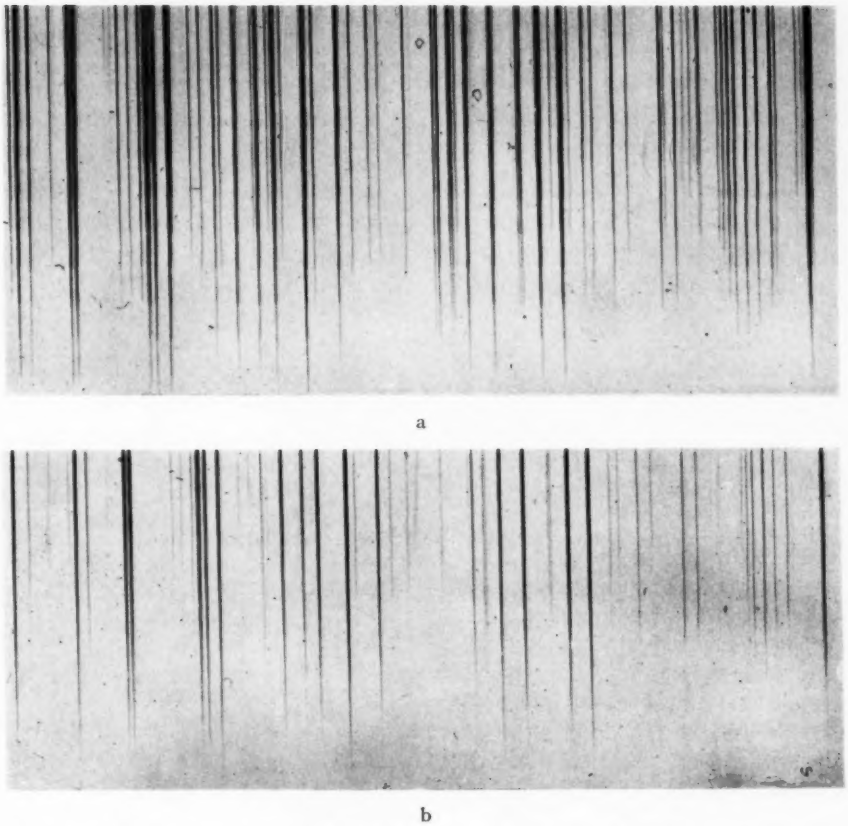


FIG. 1. (a) Typical step-spectrum taken with one of the discontinuous variation auto-calibration methods; (b) Typical wedge-spectrum taken with one of the continuous variation auto-calibration methods.

We thus use the whole course of the curve rather than a single point on it to determine the intensity ratio. Objections to this method arise from difficulties of carrying it out in practice, but when it can be used it is often more convenient, always more rapid, and sometimes more accurate, than any other. Methods of this type are listed in Table I as single exposure auto-calibration methods.

In several suggested means for varying intensity, although a single exposure may be sufficient, an auxiliary calibration source must be used. This is because the intensity variation involves some peculiarity of the calibration source, such as having a continuous spectrum or containing lines whose intensity ratios are known, which may not be a property of the source being measured.

When for one of the reasons discussed below a single exposure calibration method, cannot be used, it is necessary to make a number of separate exposures of equal duration, and this usually requires an auxiliary source unless the source of light being measured is sufficiently steady during the time required to take all the exposures necessary.

Certain calibration methods reduce the light by the same proportion at all wave-lengths; these so-called neutral methods are distinguished in Table I by being indented in the third column. Non-neutral or selective methods, while capable of giving results as accurate as neutral

methods, are less convenient because of the necessity of pre-calibration at selected wave-lengths.

In the fourth column of Table I are given the conditions which must be fulfilled in order that the method under consideration may be applicable, while the fifth column contains remarks regarding particular advantages or drawbacks of each method.

An attempt has been made to arrange the methods in Table I roughly in the order of diminishing convenience. Usually one's choice will depend on the spectroscopic equipment at hand or on the time of exposure required, since where this is very long multiple exposure methods are impracticable. When the highest photometric accuracy is required one should always use at least two systems of plate calibration and check their results against each other, since hidden errors often creep in from unexpected sources.

Three methods of determining plate response have come into common use: (1) direct measurement of density with a microphotometer,*—most direct and accurate; (2) estimation of least visible density by eye and correlation of this with intensity,—quick and convenient, but only semi-quantitative; (3) visual search for adjacent areas of equal density,—quick and accurate, but usually applicable only in cases of the type discussed in §3.

Under several of the methods listed in Table I will be found two alternatives which differ only in that one produces step-spectra as shown in Fig. 1a, and the other wedge-spectra as shown in Fig. 1b. The step-spectra are more convenient for microphotometer use since one can determine positions along the line more readily; the wedge-spectra are designed primarily for use with the limiting density methods.

§2. Discussion of Calibration Methods of Table I

(a) Auto-calibration single exposure methods

1. Step-sector or logarithmic sector disk at spectrograph slit. This method has been long and widely used on account of its convenience, but has been the subject of much controversy because it obviously does not vary the actual intensity of the light falling on the plate, but rather its average intensity. Recently, as a result of the investigations of O'Brien,[2] of Twyman,[3] and especially of Webb,[4] its use under certain conditions has been vindicated and we seem justified in placing it at the head of our list. Webb's results indicate that an intermittent exposure is equivalent photographically to a continuous exposure when the rate of flash is so great that each grain of the emulsion receives on the average not more than one quantum per flash. We realize now that the frequency of flash is the critical variable which caused such lack of correlation among the results of previous workers. Webb showed that for each intensity, emulsion and wave-length a critical frequency of flash exists above which it is quite safe to use intermittent exposures as equivalent to continuous. The critical frequency for ordinary conditions is of the order of 10 flashes per second, but to be safe one should run the sector as rapidly as possible. The greatest reciprocity and intermittency failures occur when the quanta are small (infrared), or the emulsion grains small (slow plates or papers), or the intensity high, but even under such conditions O'Brien and Parks[2] showed that with sufficiently high sector speeds the method gives accurate results.

Convenient forms of step and wedge sectors are shown in Fig. 2. One of these may be placed directly in front of the slit of a stigmatic spectrograph, mounted on the spindle of a small motor rotating at more than 1800 r.p.m. References to the use of fixed sector disks will be found

a b

Fig. 2. (a) Typical step-sector disk for high speed rotation directly in front of the slit of a stigmatic spectrograph to produce spectrograms as in Fig. 1 (a). (b) Typical logarithmic sector disk for use in producing wedge spectra as in Fig. 1 (b).

* The term microphotometer is used out of deference to common usage, although densitometer or microdensitometer are much more expressive of the function of the instrument.

in articles by Scheibe and Neuhaüsser,[5] by O'Brien and his collaborators,[2] and by Twyman and Simeon.[6] The disks can be obtained commercially.

Chief sources of error in using sector disks appear to be running the disk too slowly, and using it in a part of the beam such that all parts of the prism or grating are not simultaneously filled with light during a flash; the resolving power of the spectrograph and the shape of the spectrum lines may be changed in this way, and the intensity steps may not be truly neutral.

2. Step-weakener or wedge at slit. If one wishes to avoid even the remote possibility of errors due to reciprocity or intermittency failure, a step-weakener (or its counterpart, an optical wedge) can be used at the slit. Unfortunately no substance has been discovered which cuts light down by absorption, reflection or scattering uniformly throughout the spectrum, so that a truly neutral wedge or step-weakener does not exist. However, it is comparatively simple to measure the transmission of each step of such a weakener by means of a thermopile used in conjunction with a mercury arc and a quartz monochromator, and curves similar to Fig. 3 can be drawn and intermediate values interpolated.[7]

For the photographic infrared and the visible region one can make a moderately neutral weakener by exposing a photographic plate (preferably of fine grain such as a contrast or a positive plate) in steps of the desired width, for suitable lengths of time.[8] Both absorption and scattering produce the desired reduction of intensity here, but it is usually advisable to keep the second of these at a minimum since it varies rapidly with wave-length.

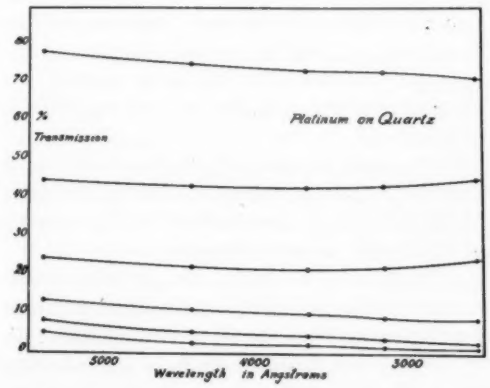


FIG. 3. Typical set of calibration curves for the various steps of a platinum-on-quartz step-weakener.

So-called neutral glasses are generally far from neutral except in limited regions of the visible, but wedges made from them can be used for photographic photometry if calibrated at suitable wave-lengths. The standard Goldberg wedge consisting of a suspension of carbon in gelatin is moderately neutral, but also over a limited range only. Thus far no material has been found more suitable for the ultraviolet than platinum or quartz, as described by Merton.[9] Either sputtering[10] or evaporation[11] can be used to make the deposit, although the characteristics of films made in the two ways differ. Where the deposited material gathers in small droplets the light is scattered rather than absorbed.

Platinum wedge and step-weakeners on quartz can be obtained commercially from Zeiss, or can be readily made in the laboratory. The platinum surface of such a weakener should be covered with another layer of quartz to protect it from dust and scratches. When this is done care must be taken that interference fringes are not formed which produce unevenness in the spectrum lines.

When any of the methods which involve placing a step or wedge reducer at the slit are used, this must of course be uniformly illuminated along its length, which process often presents difficulties. Ornstein and his collaborators[12] have described methods of achieving this result. In selecting any method to do this, it should be borne in mind that the resolving power of the spectrograph will be lowered if an image of the source be focussed on its collimator so that it is not illuminated with coherent radiation. So long as the slit and collimating lens or grating are filled with light, introduction of further optical parts will of course not increase the total light flux through the spectrograph.

In order to make sure that uneven illumination of the slit has not affected the intensity ratios it is desirable to invert the step-weakener, wedge, or sector frequently so that such effects will tend to cancel out, or at least be detected.

3. Step-diaphragms before the slit. Diaphragms (rasters) in which steps equivalent to those of a step-weakener are produced by geometrical means have been described by Hansen,[13] by Frerichs,[14] and by Miller.[15] Diaphragms having steps of various lengths are introduced into the beam, and by means of astigmatism the illumination on each part of the slit is made proportional to the length of one of the steps. As pointed out by van Cittert,[16] very large errors are sometimes introduced when these devices are assumed to be neutral, due to changing illumination of the spectrograph from step to step. Here again we are restricted to using the method with a stigmatic mounting, though Frerichs[14] has used step-diaphragms with a Rowland-type mounting by placing the diaphragm at the position of the vertical astigmatic focus outside the slit. Only a limited spectral region can be photographed at one time with this method, since the position of the step-diaphragm must be shifted for each wave-length setting.

4. Step-weakener or wedge at the plate. Concave grating mountings or certain types of prism spectrograph are frequently so astigmatic that the methods listed above cannot be used, since any point on the slit is drawn out into a vertical line in the spectrum. In the Paschen and Eagle mountings, for example, long astigmatic lines are often produced with central portions of very uniform intensity, ideally suited for superposition of intensity reducing devices. This suggests placing a reducing wedge or weakener directly in front of the plate and photographing the lines through it. This is quite practicable when narrow multiplets, Zeeman, or Stark patterns are to be measured, but one seldom has available a reducer which will cover more than two inches of plate except in the range where photographically produced reducers can be used (3500–10,000A). Experiments are at present in progress which indicate the feasibility of producing satisfactorily uniform weakeners up to 20 inches long which can be used in the range 10,000–2400A by coating platinum on Eastman AAA acetate film.

5. Theoretical astigmatism. Dieke[17] points out that one may calibrate plates by making use of the theoretical variation in intensity along a spectrum line produced by the known astigmatism of the grating. This method is especially adapted to a Paschen mounting where the grating is not touched for a long time, but depends on the uniformity of the rulings of the grating. Dieke suggests means by which the method may be useful even in unfavorable cases where the grating reflects different amounts of light from different parts of the ruling.

The other auto-calibration methods have not been developed beyond the status described in the first paper.

(b) Auxiliary source single exposure methods

1. Step-slit or wedge slit. With a stigmatic spectrograph having a fairly long slit, one can use a slit having various widths along its length, and when calibrations are then made using a source having a continuous spectrum one obtains a spectrogram consisting of several bands having known relative intensities as in the case of the step-weakener or step-diaphragm methods. This method has been successfully used by Ornstein and his collaborators[18] in the visible region and in the near ultraviolet, where an incandescent lamp can be used to give the continuous spectrum, and by Thompson and Duffendack[19] in the ultraviolet, using a hydrogen discharge tube as a continuous source.

The method is a neutral one, but has limited application, for the slit must never be so narrow that diffraction errors result, while if it is too broad overlapping of wave-lengths occurs. Of greater importance is the fact that a continuous spectrum is being used to calibrate a plate on which line spectra are being measured, and the Eberhard effect may introduce large errors. This effect can be reduced by covering the plate with a diaphragm containing broad slits spaced say 25A apart while exposing to the continuous spectrum.

In the case of the wedge-slit, it is most convenient to use a width variation of the form illustrated in Fig. 4, the width of the slit varying logarithmically along its length.

2. Theoretical line ratios. In certain simple spectra the ratios of lines in multiplets are known quite accurately from theory and experiment. In zinc, for example, certain lines have the intensity ratio 5 : 3 : 1, and it is conceivable that one might introduce a small quantity of zinc salt into

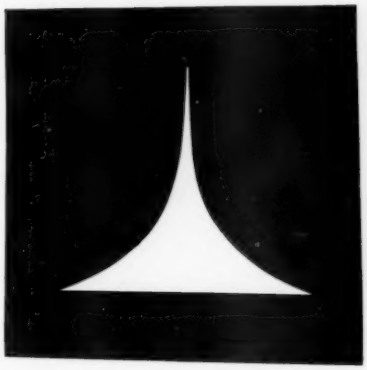


FIG. 4. Convenient shape of a wedge-slit to produce wedge-spectra with a continuous source.

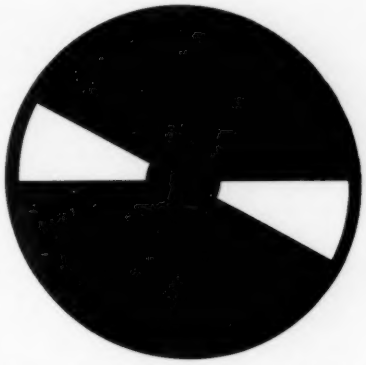


FIG. 5. Common form of radial diaphragm with variable aperture for use with prism instruments.

the source being studied if it be an arc or spark, and utilize the known ratios of these lines for calibrating the plate. This method has not been extensively tested, and should not be trusted unless the greatest care is taken to eliminate self-reversal and excitation errors.

(c) Multiple exposure methods

If none of the methods listed above is satisfactory, or if one wishes to check on the accuracy of a particular application of one of them, it is necessary to use a multiple exposure method, which, unless the source being studied is steady, involves using an auxiliary source which possesses this characteristic (and, if desired, an auxiliary spectrograph). Usually an incandescent lamp is most satisfactory for the region 10,000–3800A (or to 2800A if a quartz window is provided), a mercury arc for the region 5800–2300A, and a hydrogen discharge tube[20] for 7000–900A. Certain steady sparks can also be used for calibration.[21]

1. Variable aperture diaphragms. One of the simplest methods of intensity variation, but one usually of doubtful accuracy, is that in which the cross section of a collimated beam of light is altered by means of a variable aperture diaphragm. This may be done readily if the beam in which the aperture is placed be of uniform distribution across a section for all wave-lengths used, and if all wave-lengths be treated alike by the spectrograph. This condition is difficult to fulfill, as most lenses have zonal aberrations so that a radial diaphragm must be used, while the prism or grating usually is symmetrical about a horizontal line if at all. A diaphragm of the shape shown in Fig. 5 comes closest to fulfilling these conditions for a prism instrument, but it obviously cuts the resolving power at least in half if placed in the usual position, and such a system should be calibrated at a number of wave-lengths.

2. Wire gauze screens. These effective multiple aperture diaphragms avoid the difficulties outlined in the last paragraph and are coming into wide usage, but they may introduce diffraction errors unless handled carefully.[22] They should have their transmission factors calibrated under the conditions of use, but are usually found to be sensibly neutral when this is done.

3. Neutral filters. If one will take the trouble to measure the transmission factors of a series of (approximately) neutral filters at various wave-lengths over the spectral range desired, he has available one of the methods which has least chance of suffering from hidden errors. The filters are introduced into the beam, preferably where it has a large section, one at a time or in groups, since they are additive. Care must be taken to see that they are moderately plane, as any prism effect may cause some light to be thrown off the slit and thus introduce refraction errors.

4. The inverse square law. Here we have the most fundamental method, and one which is very useful for testing the accuracy of other methods directly, since it can always be used by introducing a sufficiently intense source which approximates a point. A matte surface, such as a

quartz disk with both sides rough ground, is placed at the normal position of the light-source being studied, and is illuminated by an intense source such as a short section of mercury arc or an incandescent lamp, placed on an optical bench at an adjustable distance from the scattering surface. Much of the light is wasted, but if sufficient enter the spectrograph to produce a satisfactory spectrum in the time of exposure required, the intensities produced are accurately proportional to the inverse square of the distance between source and scatterer. It is particularly necessary to keep out stray light, and to avoid atmospheric refraction and absorption in long paths.

5. Variable slit-width. If we have a symmetrically opening slit (not necessary in high dispersion instruments) the intensity of the spectrum will be proportional to the slit width for medium widths when we use a continuous spectrum. This method is similar to the step- and wedge-slit methods discussed above, and is subject to the same limitations.

6. Variable slit length. When using a spectrograph having much astigmatism, such as a grating in the Paschen mounting, one can vary the line intensity by varying the length of slit illuminated. This is equivalent to the preceding method, but can be used with a line spectrum.

7. Variable sector disk. The original Abney style variable sector disk has been restored to grace for photographic photometry, provided, as with the step-sector and logarithmic spiral disks discussed above, it be rotated so rapidly that each grain of the emulsion is on the average illuminated with only one quantum per flash.

8. Variable exposure time. This method, so commonly used because of its simplicity, should only be used as a check on other methods, or when the reciprocity failure of the plate is determined directly as in the fluorescent plate method of Harrison and Leighton.[23] Under no circumstances should it be trusted for use with a given emulsion just because someone else found that this emulsion obeyed the reciprocity law. The common use of the time scale in commercial plate sensitometry has no bearing on its validity for photographic photometry.

9. Other methods listed. Production of intensity variation by varying the source excitation, or by using polarizing apparatus, while of value in other methods of photometry have little value in photographic usage.

§3. Special Eye-Match Homochromatic Methods for Absorption Spectrophotometry

In order to avoid the necessity of using a microphotometer attempts have been made to develop other methods of measuring plate response. That found most successful for routine absorption work consists in matching by eye contiguous areas of equal density, and this method has been utilized by makers of commercial instruments with great success. Two spectra from the same source are photographed simultaneously and side by side, the one reduced in intensity by a known amount throughout, and the other reduced by the absorbing material. By proper modification these methods can of course be applied to reflection and scattering measurements.

The basic principle involved is illustrated in Fig. 6. The spectra are produced in pairs,

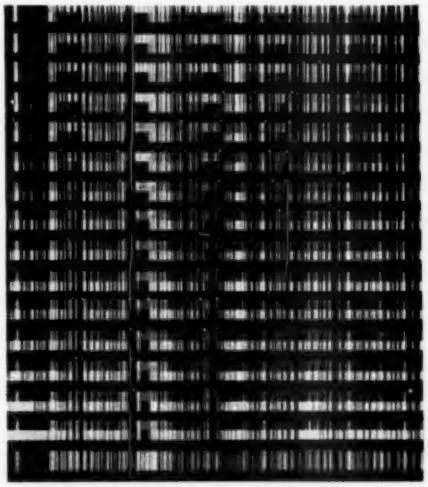


Fig. 6. Typical set of eye-match spectrograms as used in absorption spectrophotometry. (Courtesy Adam Hilger, Ltd.)

exposures to absorbed and reduced spectra being taken simultaneously with some sort of beam-splitting arrangement attached to the spectrograph. This presupposes use of a stigmatic

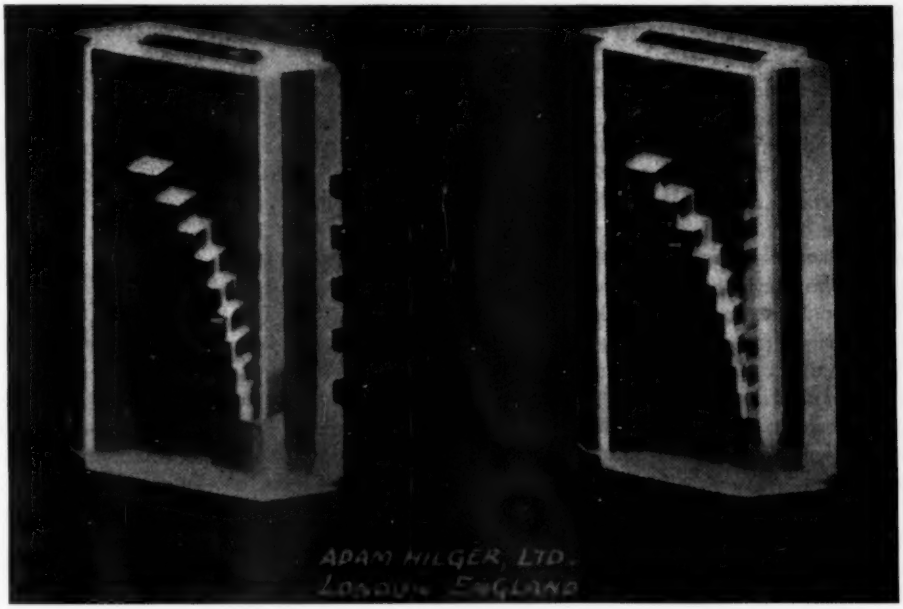


FIG. 7. Notched echelon and plain echelon cells, the first to contain the absorbing solution, and the second to contain the solvent and act as control in the Hilger notched echelon cell method of absorption photometry. (Courtesy Adam Hilger, Ltd.)

instrument, but most measurements of the absorption curves of solutions are made with quartz prism instruments of moderate dispersion in any case.

The oldest commercial device using this method is the Hilger sector photometer. This, together with the Judd Lewis sector photometer and the Jones absorption spectrophotometer, were discussed in the first paper. The old controversy about the accuracy of the Hilger sector photometer appears to be finally settled by the work of O'Brien[2] and of Webb;[3] its accurate results, delivered in spite of apparent disregard of intermittency and reciprocity laws, resulted from the high sector speeds used.

The firm of Adam Hilger, Ltd., has developed a series of successful absorption spectrophotometers which are widely distributed. The theory underlying these has been given by Twyman and others.[24] Two important new additions are the Spekker ultraviolet spectrophotometer,[25] and the notched echelon cell photometer.[26]

Various means have been utilized for varying the intensity of the reduced beams; thus the Hilger sector photometer uses a variable aperture rotating sector disk, but introduces intermittency into the absorbed beam also by a fixed-aperture rotating sector. Polarization has sometimes been used, as have variable radial diaphragms and wire screens. In the Spekker photometer the reduction in intensity is produced by a variable aperture controlled by a micrometer screw.

For the ultraviolet a steady continuous source is not always available, and indeed is not necessary with this type of apparatus where provision is made to split the beam of light from the spark or other source rich in lines into two equivalent beams. When this is done it is necessary that the two quartz rhombs or other beam-splitters shall view the spark from almost the same aspect so that variations in source position shall not alter the relative intensity of the beams.

A very ingenious method for measuring the absorption of liquids or solutions is that described by Twyman, Spencer and Harvey[27] and marketed by Hilger as the notched echelon cell. The absorption cell, of quartz, is so built as to cause the light to pass through steps of liquid of

varying thickness, so that in one exposure, spectra taken through ten different thicknesses of solution are recorded, each with an appropriate comparison spectrum. The comparison spectra are taken with a beam-splitting device, as above, and the second beam passes through a control layer of solvent and through a revolving sector producing a known and uniform diminution in intensity. The translation of plate readings into absorption curves is greatly simplified with this apparatus. The method has the advantage of simplicity, rapidity, economy of solution and of photographic materials, and increased range of utility and accuracy. The designers say that only about 0.4 çc of solution are required for obtaining a complete absorption curve, which can usually be obtained with a single exposure of 20 sec. duration. A great advantage of this type of cell is that the absorption curves of materials whose absorption is changing rapidly with time can be measured. The notched echelon cell can be used over the range 12,000–2000A.

A diagram of the notched echelon cell is shown in Fig. 7. Two cells, one with and one without notches are placed in a mount and used to throw contiguous pairs of spectra on the slit of the spectrograph. A detailed description of the method will be found in the references given.

The ability of the eye to select from a number of pairs of contiguous lines two lines whose densities are equal is very good, and the accuracy of the method is such that use of a microphotometer is unnecessary.

The third method of determining plate response mentioned above, using the least visible density, has not yet been developed to the accuracy necessary for quantitative work.

§4. Heterochromatic Photometry

In order to carry out the comparison of intensities of spectrum lines of widely differing wave-length one has recourse to heterochromatic photometry, which involves the additional step of correlating the plate calibration curves obtained for the various wave-lengths. This process is illustrated in Fig. 8, where we depict the operation of fitting the two-dimensional d-log I curves for different wave-lengths into a three-dimensional d-log $I-\lambda$ graph.

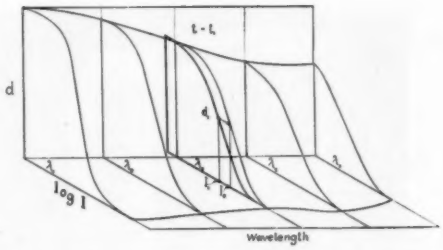


FIG. 8. A typical three-dimensional graph connecting density, log intensity, and wave-length, to illustrate the fundamental principles of heterochromatic photometry. A calibration curve for a particular wave-length is being fitted into the three-dimensional graph by means of the standardization spectrum.

The various d-log I curves are plotted on different log I scales, since all intensity variations in them are relative rather than absolute. We know λ, of course, and can measure d with a microphotometer, but we must provide some means of determining I to connect the different curves. This can be done by photographing under the standard conditions (of time, development, etc.), and of course on the same plate, a source of light whose intensity distribution with wave-length is known. Such standard sources have been difficult to obtain in the past, especially in the ultraviolet region, and this least satisfactory aspect of photographic photometry appears to have been improved less in the last few years than almost any other. The black body run at a known temperature, and the tungsten filament run at a known color temperature continue to be most widely used for the range 10,000–3000A, supplemented between 4350 and 2400A by the mercury arc run under controlled conditions.

For standardization of plates one requires a source, preferably giving a line spectrum, having constant energy emission in time, and of sufficient intensity over the desired wave-length range so that the energy distribution can be measured by means of a nonselective spectro-radiometer. In addition to the sources mentioned above one can use hydrogen[28] or helium[29] discharge tubes, while the new sodium, cadmium and other vapor lamps give promise of being useful for this purpose. Where the source constancy is not perfect a monochromatic illuminator sensibly free from scattered light, as described by van Cittert,[30] or by Forsythe and Barnes,[31] can be

used in the manner described by Dorgelo,[32] one wave-length at a time being measured and sent into the spectrograph.

In designing new systems of plate standardization it should be borne in mind that determination of the variation of plate sensitivity with wave-length is not enough, since the overall variation of spectrograph transmission factor, of the mode of focussing, and of the character of illumination is required in addition. These are all problems which can be settled by homochromatic photometry, however, so a method which gave a known plate sensitivity variation such as that proposed by Harrison and Leighton[33] would simplify the problem under many conditions.

A source of error in heterochromatic photometry which has received little attention is that of the variation of line shape with mode of spectrograph irradiation, an effect which may cause grave differences of response to different types of source. The theoretical aspects of this question have been discussed by van Cittert,[34] while Stockbarger and Burns[35] have carried out extensive investigations on the experimental side. Their work indicates the great desirability of using a standard source having the same physical characteristics of the unknown (line shape and breadth, and shape, size and position of source).

§5. The Extreme Ultraviolet

While comparatively little photographic photometry has been carried out in the photographic infrared, with the new Eastman plates[36] and certain obvious modifications of technic[37] this presents few difficulties. In the short wave-length region (2000–50A), on the other hand, the difficulties of finding suitable methods of varying light intensity in vacuum instruments has required the devising of special methods.

Harrison and Leighton[38] have described a method of homochromatic photometry using fluorescent oils which has the advantage of giving constant contrast over long regions of the spectrum. Thus one needs only one calibration curve per plate, and the fluorescent emulsion can be exposed in a vacuum spectrograph, and then calibrated in air with a quartz prism instrument. The process can often be simplified by using time variation for calibration with the vacuum instrument, and then testing the reciprocity law failure at the longer wave-length.

Ladenburg and Van Voorhis[39] have published homochromatic results obtained by using Eastman extreme ultraviolet sensitive plates (an ordinary emulsion coated with a deposit of fluorescent crystals), in which intensity variations were produced by varying the aperture of the beam with a diaphragm. While they obtained excellent results this method must be used with caution, as the common imperfections in fluorite optical parts, or the non-uniform reflections of mirrors and gratings may cause errors.

The work cited in §2 regarding the rotating sector disk makes this now applicable to the extreme ultraviolet if used directly at the slit of a vacuum spectrograph and rotated at high speed, which opens new possibilities for the short wave region. Also the recent work of O'Bryan[40] on thin deposits of metals on thin celluloid films opens the possibility of developing step-weakeners for the short wave region.

Little progress has been made in heterochromatic methods for the vacuum region, although several fluorite spectro-radiometers have been constructed for use in the range 2300–1300A.[41] A method of producing a known variation of plate sensitivity with wave-length has been used by Harrison and Leighton,[33] and this would probably find its greatest use in the short wave region. Contrary to the results of Anderson and Bird,[42] who found equal energy sensitivity over a wide range of all types of fluorescent materials tested, Harrison and Leighton found variation of sensitivity according to complex laws for most materials, but uniform quantum sensitivity for some. Such a simple law of response would give the required condition for changing heterochromatic problems into simpler homochromatic ones, since the emulsion would always be exposed to quanta of the same size no matter what the incident wave-length, and only the transmission characteristics of the spectrograph system for various wave-lengths need then be determined.

§6. Densitometers and Microphotometers

The trend in the construction of apparatus for measuring densities has been away from the large universal instruments and toward diversi-

fication for various purposes. Designers now usually have in mind some particular feature such as high sensitivity or accuracy, quick action, stability in operation, ease and convenience of manipulation, constant availability for use, or simplicity and cheapness. No longer does the recording microphotometer completely overshadow the direct reading instrument in importance as the advantage of the latter in measuring selected relative intensities of lines and bands is becoming more appreciated. For depicting and measuring line contours and patterns, or indicating the general course of intensities the recording microphotometer is of course invaluable, but for many other purposes the recording feature is unnecessary and cumbersome.

The photoelectric cell has almost displaced all other devices as sensitive receiver in microphotometers, as is indicated by the large proportion of instruments recently constructed using this.[49–58] The Moll instrument is still widely sold, however, its manufacturers having successfully increased the rapidity of action of its thermocouple-galvanometer combination to compete, to a certain extent, with photoelectric mechanisms.

The newer types of photovoltaic cell have come into some use for densitometers, instruments using them having been described by Lange[59] and by Milligan.[60] Bausch and Lomb Optical Company now manufacture a very compact instrument using the photronic cell, especially designed for simple density measurements of the type needed for quantitative spectroscopic analysis. This is in a lower price class than most other commercial densitometers.

Projection-type instruments similar to that described by Hesthal and Harrison[54] have been constructed by Beals,[55] by Duane,[56] and by Jones and Brose[57] while Harrison[58] has described a somewhat more complicated dichromatic instrument which is free from the Schwarzschild-Villiger effect and runs on low voltage batteries, while retaining the advantage of projecting a large portion of the plate being measured on a screen.

The chief disadvantage of photoelectric instruments using valve amplification is apt to be instability, drifts being set up in the deflecting instrument when the apparatus is turned on so that a preliminary warming-up period is required. In large laboratories using the instrument constantly enough so that tube filaments can be left burning day and night this is of small consequence, but it does constitute a serious defect otherwise.

An instrument having a very stable zero has been described by Smith and Wilson.[61] They use a radiometer of the Nichols type, which is more sensitive and stable than many thermoelectric instruments.

The optical systems used by most designers show few novel features, the commonest being microprojection as used in the majority of commercial instruments, with one, two, or three slits. Many of the instruments described in the literature present practically nothing new, though in some ingenious methods of coupling plate drive and record drive have been used, as in that of Weigle[62] where the galvanometer is made to rotate.

Richtmyer and Hirsh[63] and Harris[64] have discussed the question of reducing grain effects in microphotometer readings, as have Wright,[65] Linnik[66] and Kellström.[67] The Richtmyer and Hirsh device was especially designed for use with the Moll microphotometer.

The readings of microphotometers are usually recorded in terms of linear deflections of the galvanometer or electrometer used with them. In order to obtain actual line intensities it is then necessary to change these deflections into densities or blackenings,* and by means of the calibration and standardization curves of the plate change these into intensities. Wouda[68] has described an instrument for carrying out the first part of this operation rapidly. An image of a long straight-filament lamp is moved across a plot of the calibration curve by the deflection of a galvanometer mirror and the intensities can be read off directly from the intersection of the filament image and the curve. An even more useful device is that of Thompson,[69] who sends a long line of light from the galvanometer mirror

* The term blackening has been used by several investigators to denote the ratio $(I_0-I)/I_0$, where I_0 represents the microphotometer deflection through the clear plate, while I represents that through the part of the plate whose blackening is being determined.

through a template slit cut to the predetermined shape of the calibration curve for a portion of the plate, that section of this line which is transmitted through the slit being in a position indicating the intensity of the line directly. These devices are of great value where Zeeman or hyperfine structure patterns which require high resolving power are being examined, since here it is impossible to use a wide slit to produce flat-topped lines. The curves obtained can be integrated directly with a planimeter as they come from the modified microphotometer.

A still further improvement would be to make the apparatus provide its own determination of the calibration curve as needed, and to have it also automatically apply a correction for plate standarization. The details of a comparatively simple mechanism to do this have been worked out, and it is now being constructed in the writer's laboratory. Recent developments in servomechanisms make possible a number of features hitherto considered impracticable.

Bibliography

1. G. R. Harrison, J.O.S.A. and R.S.I. **19**, 267 (1929).
2. B. O'Brien and V. L. Parks, Phys. Rev. **41**, 387 (1932).
B. O'Brien and E. D. O'Brien, ibid. **37**, 471 (1931).
B. O'Brien, ibid. **36**, 381 (1930) and **33**, 640 (1929).
3. F. Twyman, various publications of Adam Hilger, Ltd.
F. Twyman and G. F. Lothian, Proc. Phys. Soc. **45**, 643 (1933).
4. J. H. Webb, J.O.S.A. **23**, 157 and 316 (1933).
5. G. Scheibe and A. Neuhaüsser, Zeits. f. Angewandte Chemie **45**, 1218 (1928).
6. F. Twyman and F. Simeon, Trans. Opt. Soc. (London) **31**, 169 (1930).
7. Cf. H. Kienle, Zeits. f. Astrophysik **1**, 13 (1930).
8. Cf. G. Haase, Zeits. f. Instrumentenk. **52**, 529 (1932).
9. T. R. Merton, Proc. Roy. Soc. **A106**, 378 (1924).
10. B. O'Brien, Proc. Opt. Soc. Am., Buck Hill Falls Meeting, Oct. (1933). J.O.S.A. **24**, Feb. (1934).
11. R. C. Williams and G. B. Sabine, Astrophys. J. **77**, 316 (1933).
12. Ornstein, Moll and Burger, *Objektive Spektralphotometrie*, Vieweg, 1932, p. 83.
13. G. Hansen, Zeits. f. Physik **29**, 356 (1924), Ann. d. Physik **383**, 558 (1925), A. v. Hippel, ibid. **80**, 672 (1926), K. Schachtschabel, ibid. **81**, 929 (1926).
14. R. Frerichs, Ann. d. Physik **386**, 809 (1925).
15. O. E. Miller, Rev. Sci. Inst. **3**, 30 (1932).
16. P. H. van Cittert, Zeits. f. Physik **73**, 249 (1931).
17. G. H. Dieke, J.O.S.A. **23**, 274 (1933).
18. Reference 12, p. 86.
19. K. B. Thompson and O. S. Duffendack, J.O.S.A. **23**, 101 (1933).
20. See for example G. B. Kistiakowsky, Rev. Sci. Inst. **2**, 549 (1931) or Bay and Steiner, Zeits. f. Physik **59**, 48 (1929), ibid. **45**, 337 (1927).
21. A. Hollaender and J. P. Foerst, Rev. Sci. Inst. **4**, 347 (1933) and references therein; also G. S. Forbes and F. P. Brackett, J. Am. Chem. Soc. **53**, 3973 (1931).
22. G. R. Harrison, J.O.S.A. and R.S.I. **18**, 492 (1929); cf. also C. W. W. Read and L. W. Johnson, Phil. Mag. **11**, 1152 (1931).
23. G. R. Harrison and P. A. Leighton, J.O.S.A. **20**, 313 (1930).
24. F. Twyman and F. Simeon, Trans. Opt. Soc. (London) **31**, 174 (1929–30) and Hilger publications.
25. F. Twyman, ibid. **32** (1930–31); also Hilger publication No. 184.
26. F. Twyman, Proc. Phys. Soc. **45**, 1 (1933).
27. F. Twyman, L. J. Spencer and A. Harvey, Trans. Opt. Soc. (London) **33** (1931–32) and Hilger publication No. 178.
28. For calibration methods by means of filters, see D. C. Stockbarger and L. Burns, Phys. Rev. **34**, 1263 (1929).
29. L. S. Ornstein and H. C. Burger, Zeits. f. Physik **76**, 777 (1932).
30. P. H. van Cittert, Physica **3**, 182 (1923); Rev. d'Optique **2**, 57 (1923).
31. W. E. Forsythe and B. T. Barnes, Rev. Sci. Inst. **4**, 289 (1933).
32. H. B. Dorgelo, Phys. Zeits. **26**, 771 (1925).
33. G. R. Harrison and P. A. Leighton, Phys. Rev. **38**, 899 (1931).
34. P. H. van Cittert, Zeits. f. Physik **65**, 547 (1930).
H. C. Burger and P. H. van Cittert, Zeits. f. Physik **81**, 428 (1933).
35. D. C. Stockbarger and L. Burns, J.O.S.A. **23**, 379 (1933) and Phys. Rev. **37**, 920 (1931).
36. *Photographic Plates for Use in Spectroscopy and Astronomy*, Eastman Kodak Co., 1933.
37. A. L. Schoen, J.O.S.A. **14**, 179 (1927).
38. G. R. Harrison and P. A. Leighton, J.O.S.A. **20**, 313 (1930).
39. R. Ladenburg and C. C. Van Voorhis, Phys. Rev. **43**, 315 (1933).
40. H. M. O'Bryan, J.O.S.A. **22**, 739 (1932) and Phys. Rev. **44** (1933).
41. P. A. Leighton, private communication, and G. R. Harrison, unpublished.
42. W. T. Anderson, Jr. and L. F. Bird, Phys. Rev. **32**, 293 (1928).
43. Symposium on photo-cells, Trans. Opt. Soc. (London) **31**, 233 (1929–30).
44. J. A. Carroll and E. B. Moss, M. N. Roy. Ast. Soc. **91**, 191 (1930); E. Gambetta, Comptes Rendus **196**, 906 (1933); A. G. Winn, Trans. Faraday Soc. **29**, 689 (1933).

45. F. C. Toy, J. Sci. Inst. **7**, 253 (1930).
46. H. G. Heil, Phil. Mag. **11**, 736 (1931).
47. D. Chalonge and P. Lambert, Rev. d'Optique **10**, 405 (1931).
48. J. K. Lees, J. Sci. Inst. **8**, 273 (1931).
49. J. and J. F. Thouvert, J. de Physique **7**, 475 (1930); ibid. **7**, 245 (1932); ibid. **3**, 1235 (1932).
50. W. Schütz, Phys. Zeits. **34**, 566 (1933).
51. G. Todesco, N. Cimento **9**, 138 (1932).
52. W. T. Astbury, Proc. Roy. Soc. **A123**, 575 (1929), and J. Sci. Inst. **6**, 192 and 207 (1929).
53. G. A. Boutry, Comptes Rendus **196**, 1101 and 1344 (1933).
54. G. R. Harrison, J.O.S.A. **19**, 301 (1929).
55. C. S. Beals, M. N. Roy. Ast. Soc. **92**, 196 (1932).
56. W. Duane, Phys. Rev. **40**, 132 (1932).
57. E. Gwynne Jones and H. L. Brose, J. Sci. Inst. **8**, 145 (1931).
58. G. R. Harrison, Rev. Sci. Inst. **3**, 572 (1932).
59. Lange, Zeits. f. tech. Physik **13**, 600 (1932).
60. W. O. Milligan, Rev. Sci. Inst. **4**, 496 (1933).
61. S. Smith and D. C. Wilson, Jr., Astrophys. J. **76**, 117 (1932).
62. J. Weigle, Rev. Sci. Inst. **4**, 595 (1933).
63. F. K. Richtmyer and F. R. Hirsh, Jr., Rev. Sci. Inst. **4**, 353 (1933).
64. S. Harris, Rev. Sci. Inst. **4**, 598 (1933).
65. F. E. Wright, J.O.S.A. **21**, 485 (1931).
66. W. Linnik, Zeits. f. Physik **61**, 700 (1930).
67. G. Kellström, Roy. Soc. Sci. Upsala **38**, May (1932).
68. J. Wouda, Zeits. f. Physik **79**, 511 (1932).
69. N. Thompson, Proc. Phys. Soc. **45**, 441 (1933).

The Development of the Photographic Objective[1]

R. KINGSLAKE, *Institute of Applied Optics, University of Rochester, Rochester, N. Y.*

(Received January 10, 1934)

INTRODUCTION

IN this paper, an attempt will be made to follow the development of photographic lenses from 1812 to the present day, chiefly stressing the reasons which led designers to adopt the types they did, and trying to follow their strivings after something better. The period to be discussed divides itself naturally into two parts, the "old" period from 1812 to 1886, and then the "anastigmat" period from 1886 to the present day. The separation of these periods was brought about by the introduction of the barium glasses by Abbe and Schott at that time. The discussion will be confined to the most prominent lens types, rather than attempting to list every lens made.

THE LANDSCAPE LENS

The first camera was merely a camera obscura with a photographic plate substituted for the ground glass screen of the earlier instrument. The lens in the camera obscura was originally a simple biconvex crown-glass lens, which would give fair definition in the center of the picture, but the image rapidly deteriorated at points more than a few degrees out. The first real attempt to improve this lens was made in 1812 by W. H. Wollaston, who suggested the use of a simple meniscus lens having a stop or diaphragm in front, with the concave side of the lens facing the diaphragm. In this way he produced his "Periscopic" lens giving quite good pictures at an aperture of $f/8$, and excellent ones at $f/16$, the definition being quite sharp out to as much as 25° from the axis. This lens has remained the most used type of photographic lens up to the present day, being fitted to millions of the world's cheapest cameras.

The first attempt to improve the simple Wollaston meniscus lens was made by C. Chevalier in 1821. He retained the meniscus outward form, but managed to achromatize the lens by the use of a negative flint and positive crown lens cemented together. In this way he removed the axial and transverse chromatic aberrations, making a lens in which the visual and actinic light came to the same focus and in which the images in different colors were of the same size. To appreciate the value of this improvement, we should realize that with a simple lens the difference between the visual and actinic focus varies with the distance of the object, so that no universally satisfactory means of allowing for it is available except in "fixed focus" cameras such as are used today. Further developments were made by T. Grubb in 1857 who varied the construction by placing the crown lens in front, thus having all three faces concave towards the stop; and by J. H. Dallmeyer who in 1865 suggested splitting the crown component into two, placing one on each side of the flint lens. In this way he was able to cover a field of 37° from the axis. This represented the limit to which these simple types of landscape lens could be carried, before the introduction of the "new" glasses.

THE PORTRAIT LENS

In the early days of the Daguerrotype, a simple landscape lens working at $f/11$ or $f/16$ could be made to give useful photographs of extended objects outdoors, by allowing a sufficiently long exposure, but it was of far too small an aperture for easy portraiture indoors. The story goes that a Professor Ettingshausen saw the Daguerrotype process in operation in Paris in 1840, and realized what an advantage it would be to have a lens of much larger relative aperture. He therefore approached Professor J. Petzval of Vienna, a mathematician only 33 years old, suggesting that he should try to design such a lens. Within about a year, Petzval had

[1] The substance of an Invited Paper given at the Annual Meeting of the Optical Society of America, October, 1933.

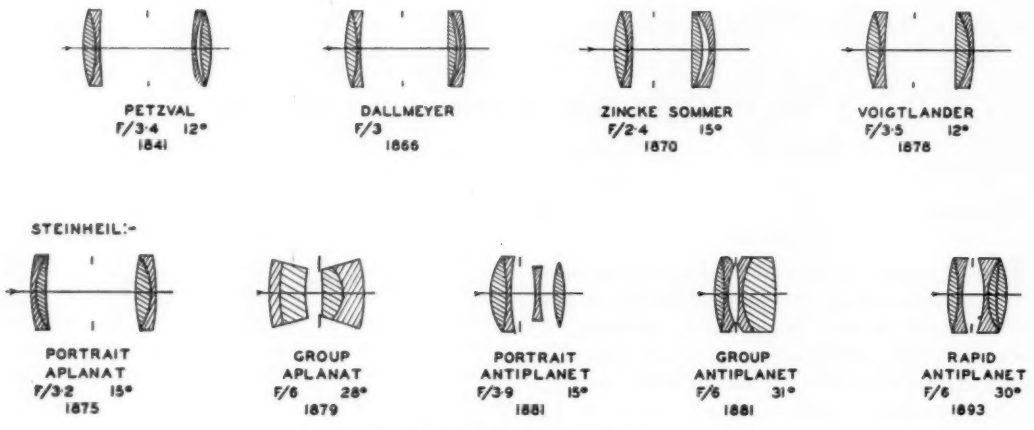


FIG. 1. Some portrait lens types.

produced the required design for a lens of aperture $f/3.4$, giving good definition over a field of 10° or 12° from the axis. This was amply large enough for ordinary portraiture, as it is in fact a real advantage if the surrounding objects are not quite so sharply focussed as the subject himself. The lens was made by the firm of Voigtländer of Brunswick.

This portrait lens of Petzval (Fig. 1) was an amazing accomplishment. Unfortunately we do not know by what methods he worked, but whatever they were, he had to invent and develop his procedure as he went along. He had no hint whatever from other lenses, and the only glasses available were ordinary crown and flint. It is not too much to say that his lens made portrait photography possible, and its popularity has been such that it is still manufactured and used in considerable numbers; it was the fastest lens made until about 1910 when anastigmats of greater aperture became available.

The original lens of Petzval was improved in subsequent years by other designers, notably by J. H. Dallmeyer in 1866 (Fig. 1), who inverted the order of the separated components of the rear element, thus increasing the aperture somewhat; by H. Zincke Sommer in 1870 who raised the aperture to $f/2.4$ and the semi-field to 15°; and lastly by Voigtländer in 1878 who contrived to cement the back element as well as the front, thus removing two glass-air surfaces. R. Steinheil in 1875–1893 designed and made a number of interesting lenses under the names of "aplanats" and "antiplanets," for portrait and group photography, in which he tried to flatten the field by a partial fulfillment of the Petzval theorem (see below). These lenses were introduced too late, and were superseded by the anastigmats before their merits were recognized.

Effects of Glass-Air Surfaces

Each glass-air surface in a lens reflects back some 5 percent of all the light that falls upon it. This makes some diminution of the brightness of the picture, but its most serious effect occurs when this once-reflected light is again reflected, this time back into the camera. Consequently a lens having many glass-air surfaces sends much stray light into the camera, tending to illuminate the shadows and give a flat-looking, non-contrasty picture. Sometimes, too, these repeated internal reflections give rise to real images of distant objects, or of the iris diaphragm, which may happen to come sufficiently near the photographic plate to form a "flare spot" or ghost image. This scattering of light by glass-air surfaces accounts for the high degree of brightness and contrast noticeable in photographs taken with a camera equipped with a simple landscape lens, as compared with the same photograph taken with a complex anastigmat.

Theoretical Advances

Although Gauss in 1841 had developed the paraxial theory of lenses very completely by his discussion of principal planes, no systematic

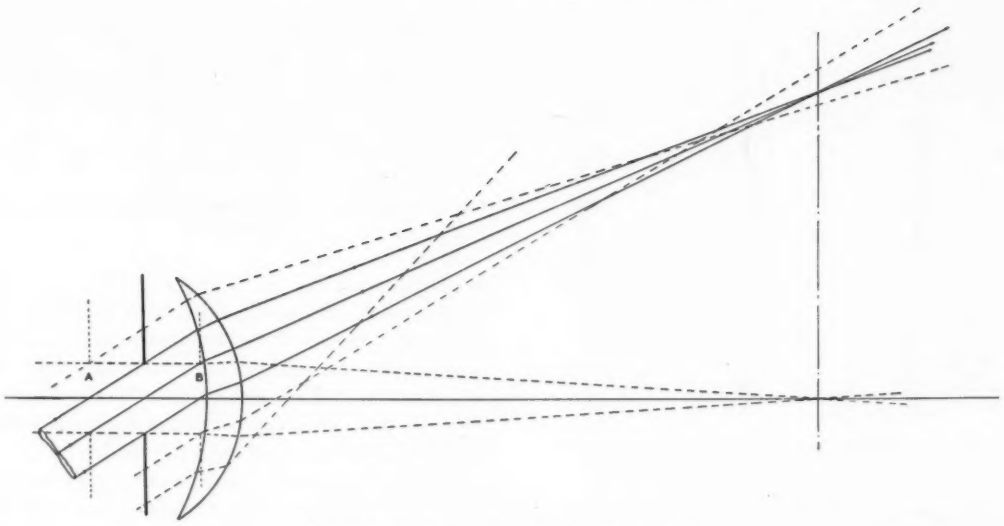


FIG. 2. Path of rays through a landscape lens.

study of the aberrations had been made until 1855 when Seidel published a number of epoch-making papers on the theory of lenses. He recognized five separate "aberrations," now well known as spherical aberration, coma, astigmatism, curvature of field, and distortion. Seidel's formulae[2] for the magnitude of these aberrations in a given case are still used, although sometimes in a somewhat modified form, by designers today.

One of the most fundamental and useful laws arising from the general theory of lenses is that giving the effects of shifting the stop along the axis, on the magnitude of the aberrations. It is clear from the diagram (Fig. 2) that the presence of the stop serves to isolate a narrow bundle of rays out of the entire incident beam, and that a stop at a different distance from the lens would isolate a quite different bundle, coming to a focus of a different sort and at a different position. For example, a stop at the position marked A in Fig. 2 would isolate a bundle of which the upper and lower rays cross at a point decidedly below the central ray, whereas if the stop were at B, the upper and lower rays would cross above the central ray. Hence if the stop is close to the lens, the image has large negative coma, and if it is too far from the lens, the coma is positive. It is to be expected that at one particular position, the coma will be entirely corrected, which is indeed the case. It should be noticed too that this stop position for zero coma also isolates those rays which will give a field turned as far away from the lens as possible.

The algebraic statement[3] of the effect of stop-shifts on the various aberrations is given by these equations:

$$
\begin{aligned}
Sph^* &= Sph, \\
Coma^* &= Coma + K \cdot Sph, \\
Ast^* &= Ast + 2K \cdot Coma + K^2 \cdot Sph, \\
Ptz^* &= Ptz, \\
Dist^* &= Dist + \tfrac{1}{2}K \cdot Ptz + 3K \cdot Ast \\
&\qquad + 3K^2 \cdot Coma + K^3 \cdot Sph.
\end{aligned}
$$

Here the asterisk indicates the changed value of an aberration due to a stop shift defined by the quantity K. Thus we notice that if a lens has some spherical aberration as well as some coma, the coma may be removed by a suitable value of K, i.e., by a suitable choice of stop position. We can now appreciate the process of designing a

[2] M. Born, *Optik.* (Springer), page 101.

[3] A. E. Conrady, *Applied Optics and Optical Design,* (Oxford), page 343.

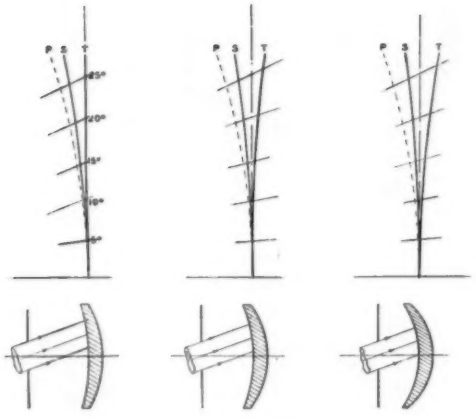


FIG. 3. Effect of bending a simple landscape lens.

landscape lens. Such a lens must have some spherical aberration in it, in order that its coma (a far more serious defect in a photographic lens) may be removed by a suitable choice of stop position. Having removed the coma, rays are traced through this stop to determine the curvature of field. Then the shape, or "bending," of the lens is changed and the process repeated until the curvature of field is what is required. In Fig. 3 three shapes of a simple Wollaston meniscus lens are shown in each of which the stop has been placed at the correct position for the elimination of coma, and the shapes of the sagittal and tangential fields are shown. (The "sagittal field" is the surface containing the foci of radial lines in the image; the "tangential field" contains the foci of lines lying tangential to the field.) It should be noticed that in Fig. 3 there is a fixed dotted curve shown in each diagram, and that the tangential image at any given obliquity is always about three times as far from this dotted curve as the sagittal image. This dotted curve is the "Petzval surface" and represents the curvature of the field as computed by paraxial formulae, in the absence of any astigmatism. It has a very constant curvature, depending only on the structure of the lens, and not at all on the distance of the object or on the stop position. It is represented by the symbol *Ptz* in the list of stop-shift effects on page 75. The possible modes of distribution of the astigmatic focal lines relative to this Petzval surface are shown in Fig.

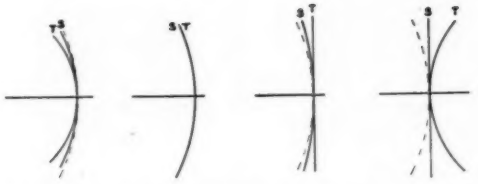


FIG. 4. Astigmatism. Relation between Petzval surface and image curves.

4, corresponding to various values of the astigmatism, negative, zero, and positive, respectively. Hence if the Petzval field is curved, there must be either curvature of field or astigmatism or both present; but if the Petzval field is flat, then zero astigmatism is automatically accompanied by a flat field.

Since with simple thin lenses this Petzval surface is inevitable, the designer has to decide how he will compromise between curvature of field and astigmatism in designing his lens. A favorite choice is to have as flat a tangential field as possible; but many lenses are made in which other compromises have been adopted.

DISTORTIONLESS LENSES

About the year 1858, the wet collodion process of photography had become definitely established, thus making indoor architectural photography possible with landscape lenses; then for the first time the inevitable distortion of those lenses became sufficiently marked to be objectionable. J. T. Goddard, in 1859, therefore attempted to design some lenses in which distortion was considerably reduced. The first, his "Double Periscopic" consisted of the simple meniscus lens of Wollaston's "Periscopic" type together with a zero-power meniscus-shaped doublet placed between the stop and the lens. This design was subsequently improved by T. Dallmeyer in 1888 in his "Rectilinear Landscape" lens. Other workers suggested various distortionless lenses (Fig. 5), which soon tended to become symmetrical about a central stop. There are so many advantages in the use of a symmetrical construction as to deserve a special study.

SYMMETRICAL LENSES

About 1860 it gradually began to be realized that a lens symmetrical about a central stop is

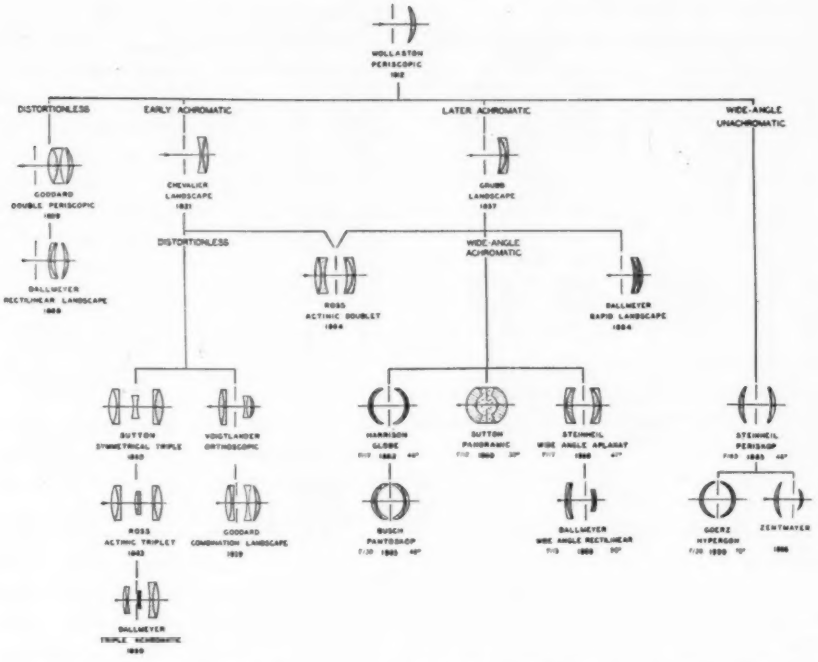


FIG. 5. Early landscape and wide-angle types. (Not corrected for spherical aberration.)

automatically corrected for distortion, merely by virtue of its symmetrical construction. Moreover, such a lens is automatically freed also from the other transverse aberrations, coma and transverse chromatic aberration. A deeper analysis by R. H. Bow and T. Sutton showed that these corrections are not complete except for unit magnification, but that distortion and transverse chromatic aberration are corrected if the system is spherically and chromatically corrected relative to the entrance and exit pupils. But even if this condition is not perfectly satisfied, these three transverse aberrations are still greatly reduced in magnitude.

Symmetrical lenses had been made by T. Davidson in 1841, who placed two Chevalier landscape lenses face to face about a central stop, and by G. Cundell in 1844 who did the same thing with two Wollaston periscopic lenses. Neither of these men, however, appears to have realized the advantages of the symmetrical construction and the double lenses never became popular.

From 1860 to 1866, a craze came in for making lenses to cover a field of an extremely wide angle. Sutton in 1860 invented his "Panoramic" lens, containing water, with a butterfly diaphragm to equalize the illumination over the field. This lens covered a field of 30° from the axis, at an aperture of $f/12$. In 1862, Harrison and Schnitzer of New York produced the "Globe" lens, so called because its outer surfaces formed part of a single sphere. This was very popular for a considerable time, and covered a field of 46° from the axis at $F/17$. This lens was later modified by Busch in 1865 as the "Pantoskop," covering 49° at $F/30$. In 1865, A. Steinheil produced the "Periskop," consisting, like Cundell's lens, of two identical simple Wollaston menisci face to face. This covered a 46° field at $f/40$, but suffered from a difference between actinic and visual foci, since it was unachromatized. Zentmayer, in Philadelphia, suggested making a hemisymmetrical doublet like the Periskop, in which the rear element was a small-scale model of the front element. A real advance was made by T. Ross in

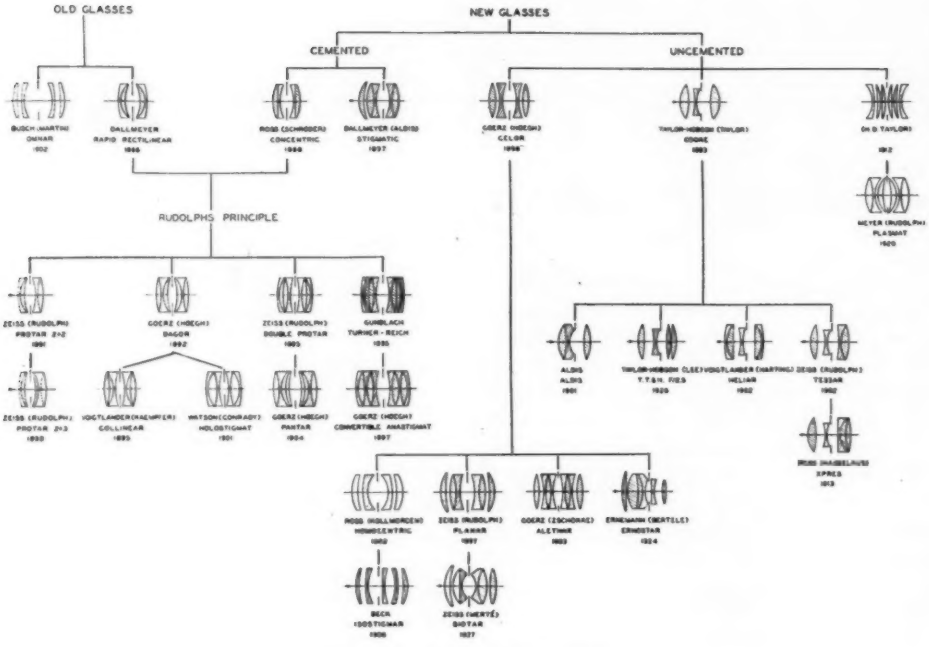


FIG. 6. The chief anastigmat types.

1864, who combined together two achromats, one of the Chevalier type and the other of the Grubb type, to make the "Actinic Doublet." A diagram of this lens suggests the later "Protar" of Zeiss, but it was of quite dissimilar construction. The climax of the wide-angle craze was reached by Goerz in 1900 with the unachromatized "Hypergon," which covers a field of no less than 70° from the axis at an aperture of $f/30$.

By far the most successful symmetrical lens was the "Rapid Rectilinear" which was introduced simultaneously and independently by Dallmeyer and Steinheil in 1866. In appearance, this resembled two identical Grubb achromatic landscape lenses facing a central stop (Fig. 6), but the construction was in reality quite different. If we refer again to the formulae giving the effects of a shift in the stop position on the aberrations, we see that a suitable choice of stop position may be used to remove astigmatism if the lens has some coma, even if it is spherically corrected. By then arranging two such lenses symmetrically about a central stop, coma is automatically removed, no matter how much coma is present in the separate components. The components of a Rapid Rectilinear lens, therefore, are spherically corrected, and have large coma which is made use of to remove astigmatism by a suitable choice of stop position. Hence, by putting two such lenses together, we get a symmetrical system, which is corrected for spherical and chromatic aberration in each element; for coma, transverse chromatic aberration, and distortion by virtue of the symmetry; and for astigmatism or any desired compromise between astigmatism and Petzval curvature by a suitable separation between lens and stop. This lens, commonly called an R.R., worked at $f/8$ or $f/7$ and covered a field of some 25° from the axis very well. The only difference between Dallmeyer's R.R. and Steinheil's Aplanat was the choice of glass, as Dallmeyer used ordinary crown and flint, whereas Steinheil used light and dense flint glass. It was made by almost every manufacturer and has been an immensely popular lens for 40 years; its manufacture was only reluctantly given up about the time of the great war. The separate elements of an R.R.

cannot be used alone because of the bad coma, but the whole system was a huge improvement over all previous lenses, except portrait lenses, in that it was spherically corrected and consequently could be made to work up to $f/8$. It had only four glass-air surfaces, and the whole system was sufficiently compact to reduce the vignetting effect of the barrel to a reasonable minimum. The R.R. lens completely superseded all the existing types with the exception of the Portrait Lens and the simple and achromatic landscape lenses.

The Petzval Theorem

In order to understand the nature of the complete revolution in lens designing which occurred when Abbe and Schott introduced the new barium glasses in 1886, we must consider in some detail the Petzval surface which has already been mentioned several times.

If a lens is used to form an image of an extended plane object perpendicular to the lens axis, it is evident that the middle of the object will be closer to the lens than the outer parts of the object. Hence we should expect the outer parts of the image to be decidedly closer to the lens than the middle of the image. This is the cause of the curvature of field produced by a lens. The radius of this curvature,[4] in the *central* part of the field, assuming a lens of *small* aperture, is given by R where $(1/R)=\sum[(N'-N)/rNN']$. In this paraxial formula, r represents the radius of curvature of a refracting surface, in the lens, separating media of index N and N', respectively, the summation being made for all the surfaces of the complete lens system. Thus the radius of curvature of the field is independent of the separations of the surfaces, the object distance, and the stop position. This theorem is commonly ascribed to Petzval, but it was known much before his time being stated and proved in Coddington's *Optics* which was published in 1829.

Previous to 1886, the only glass types available were those obtained by adding successively increasing amounts of lead to ordinary crown glass. Table I shows some typical "old" glasses.

Table I.

Type	Index N	Abbe number V
Hard crown	1.5175	60.5
Extra light flint	1.5290	51.6
Light flint	1.5427	47.5
Light flint	1.5746	41.4
Dense flint	1.6041	37.8
Very dense flint	1.6501	33.6
Very dense flint	1.7402	28.4

If we attempt to find practical methods of satisfying the Petzval condition, three interesting cases arise:

(a) *A single lens.* In this case, the Petzval condition becomes $r_1=r_2$. Thus the lens will only have any positive power if it is made thick and of a meniscus form. This explains the predominance of thick meniscus lenses in photographic objectives, notably in Steinheil's "Group Aplanat."

(b) *A separated achromat.* If it is required to fulfil the conditions both for achromatism and for the Petzval sum, it is necessary to have a separated combination of a convex crown and a concave flint lens, the separation between them being chosen so as to satisfy this condition.

$$(V/N)_a(1-m)^2=(V/N)_b.$$

Here suffixes a and b refer to the crown and flint lenses, respectively, and m is the ratio of the separation to the focal length of the crown lens (a). For example, by using ordinary hard crown and dense flint glasses, $(V/N)_a=39.5$, and $(V/N)_b=22.2$; thus a value of $m=0.27$ will serve to satisfy both the required conditions. This method of making an achromatic lens which also fulfils the Petzval condition seems to have been overlooked until K. Martin, of the Busch Company, designed his "Omnar" lens in 1902 based on this principle.

(c) *A cemented achromat.* Here the condition is that $(V/N)_a=(V/N)_b$, since m in the last paragraph is zero. This condition requires that the V's and N's of the two glasses should rise and fall together. A glance at the list of old glasses given above shows that the use of lead in the glass makes the N rise and the V fall, which is opposite to the requirements of the Petzval theorem. Consequently Abbe and Schott in 1880 tried to produce glasses of high index and high V, which could be used with a light flint of low index and

[4] L. C. Martin, *Applied Optics*, Vol. I, page 139.

relatively low V, to produce a thin achromat having a flat field. Their efforts were successful in 1886 when they discovered the barium crown glasses, which have just the desired property.

To show the extent of their success, it is of interest to compare the two pairs of glasses given in Table II. It is clear how much more nearly

TABLE II.

	Old achromat				*New achromat*		
	N	V	V/N		N	V	V/N
ELF	1.5290	51.6	33.8	DBC	1.6234	56.3	34.7
LF	1.5632	42.9	27.4	LF	1.5427	47.5	30.8
	(diff.)	8.7	6.4		(diff.)	8.8	3.9

alike are the values of V/N for the new glass pair as compared with old. It should be noted also how the V and N rise and fall together in the new achromat combination.

RUDOLPH'S PRINCIPLE

As soon as the new glasses were available, four or five designers at once attempted to use them in lenses. H. S. Schröder at the firm of Ross, in 1888, made a kind of Rapid Rectilinear lens with the new glasses, which was issued under the name of "Concentric," but it was not remarkably successful and was later replaced by the "Homocentric." Miethe, in the same year, designed a lens using the phosphate and borate glasses which Schott introduced at the same time as the barium glasses, but they were unstable and were later withdrawn. Mittenzwei also designed a cemented triple lens with an aperture of $f/3.4$ for portraiture; but none of these early attempts was particularly successful.

The first really successful design using the new glasses as a means of flattening the field was made by Dr. P. Rudolph, of the Zeiss Company, in 1890. He found that by making a kind of unsymmetrical Rapid Rectilinear lens, using old glasses for the front element and new glasses for the rear element, and by making the front element of zero focal power, he could make the astigmatism and Petzval curvature of the front element equal and opposite to those arising at the rear element, at the same time correcting the various other important aberrations. In the front element the contact face was dispersive, and in the rear element, owing to the abnormal order of the refractive index, it was collective; he therefore called this the "Principle of the Opposed Gradation of Refractive Indices."

A number of series of these lenses were made, of aperture from $f/4.5$ to $f/18$, covering fields of from 20° to 45° from the axis. The rear element was either a doublet for low apertures or a triplet for the larger apertures. The definition was excellent over the entire picture, and the field extremely flat. This lens was originally called the "Anastigmat," but the name was later changed to "Protar" when the word anastigmat became adopted as the generic term for all lenses having a reduced value of the Petzval sum. It is still made, especially as a wide-angle lens of about $f/12$ aperture.

In 1891, Dr. Rudolph tried to combine the normal and anomalous pairs of glasses into a single unit, by using three lenses cemented together, the central one being concave and the others convex, such that the index steadily advanced from lens to lens, and the V-value was least for the central lens. An imaginary line through the middle lens would divide the system into two parts, the front part being an old achromat and the rear part a new achromat. In this way he produced the "Triple Protar," which was however improved in 1902 as the "Ortho Protar" by using a strong biconvex central lens with a weak negative lens cemented on each side of it, making a meniscus outer form. Here again the index steadily increased from one end of the system to the other. In 1894, Rudolph produced the "Quadruple Protar" in which an old and a new achromat were cemented directly together, with their inner glasses being now quite dissimilar thus making a wider field possible with better corrections. This lens was convertible, meaning that each half could be used alone, or combined with another lens of the same type but built to a larger or smaller scale. In order that the separate units of a symmetrical lens may be used alone, each element must be independently corrected for coma and transverse chromatic aberration and distortion, since the symmetrical principle cannot now be invoked to correct those transverse aberrations. Thus the problem of designing a convertible lens is much harder than

that of designing a symmetrical nonconvertible lens such as the Rapid Rectilinear, the Celor, or the Planar.

Von Höegh of the firm of Goerz followed closely on Rudolph's footsteps in designing lenses of these types, producing in 1892 the "Dagor" which is extremely similar to the "Triple Protar" of Rudolph. He also designed in 1897 a quintuple cemented convertible combination which was soon abandoned on account of the expense of its construction. He also developed a quadruple cemented system in 1904 under the name of "Pantar," in which Rudolph's principle was used to give the required corrections.

The convertible lens reached its peak of popularity between about 1890 and 1905. The success of this type was really due to the invention of the dry plate in 1878. The photographer using the old "wet" plates had to carry such a huge amount of equipment when out on a photographic expedition, that a few extra lenses made no appreciable addition to his load, but when the much simpler field cameras with long bellows, and dry plates, became popular, the convertible "set" of lenses with a maximum aperture of $f/6$ (and later $f/4.5$) was ideal for such a camera. However, by about 1900, the hand camera, and especially the roll-film camera, had become so popular as to be almost universally used, and the covertible lens could then no longer be employed, as there was not sufficient bellows extension available and frequently no ground glass screen for focusing. It has therefore almost disappeared from the market.

The Celor Type and Its Modifications

Turning back once again to the discussion of the Petzval theorem given above, it was there shown that this awkward condition can be satisfied by suitably separating the components of an achromatic lens. Such a separated achromat suffers from bad transverse chromatic aberration, but by arranging two similar lenses about a central stop, this is automatically eliminated by the symmetrical principle and a good anastigmat can be produced. Von Höegh's "Celor" of 1898 was of this type, a barium crown being used with a light flint, in order to reduce the separation as much as possible. The process of the design is similar to that of the Rapid Rectilinear lens, but a new degree of freedom is available as we now have two separate lenses each of which may be "bent" independently of the other. By departing from strict symmetry, improvements can be made, and von Höegh in 1907 managed to raise the aperture of the Celor successfully to $f/3.5$. Several other designs of this type were made by Zschokke and Urban, of the Goerz Company; and the same general type was adopted by Steinheil in his "Unofocal" (1901) which was a symmetrical lens with a rather wide airspace between the lenses in each element. It is so called because all four component lenses have equal focal length and the same refractive index but different dispersions. The "Homocentric" of Ross, designed by Kollmorgen, was issued in 1902, which had the same general form as the Celor but all the lenses were menisci; and the Beck "Isostigmar" of 1906 had no less than five separate single lenses, the fifth being an additional negative lens placed between the other two. The presence of ten glass-air reflecting surfaces is a very real disadvantage in this type of design. K. Martin, of the Busch company, produced the Omnar in 1902, using only old glasses.

Other important modifications of the simple Celor type have been made. The chief of these is Rudolph's Planar of 1897, which was a symmetrical lens like the Celor but in which each negative lens was made into a hyperchromatic cemented negative doublet. This increased the negative chromatic aberration of the negative lenses, and was equivalent to making them of a much more dispersing glass. This lens covered a semifield of 25° at an aperture of $f/3.3$, and has been made by many other firms since. By making the Planar decidedly unsymmetrical, the Zeiss "Biotar" with an aperture of $f/1.4$ for 16 mm motion pictures, was produced in 1927.

Another remarkable modification of the Celor was the "Ernostar," made by the Ernemann Company in 1924. This lens successfully covered a picture $2\frac{1}{4} \times 3\frac{1}{4}$ inches in size at an aperture of $f/1.8$, a performance surpassing anything up to that time. The first negative lens was made into a thick cemented triplet, and the lens was entirely unsymmetrical (Fig. 6).

One of the most outstanding types of anastigmats ever produced was the "Cooke" lens designed by H. Dennis Taylor in 1893. This looks like a Celor in which the two negative lenses have been combined into a single lens, but actually a Cooke lens is designed directly from first principles and is not a modified symmetrical lens. The three lenses in it have very little power if placed into close contact, fulfilment of the Petzval condition being thus secured, the focal power being obtained by separating the elements suitably. The simplicity and cheapness of the Cooke lens has attracted many manufacturers, but the chief maker is still the original one, namely Taylor, Taylor and Hobson of Leicester, England. In 1926 Lee attained an aperture of $f/2.5$ by splitting the rear convex lens into two closely placed thin convex lenses.

Other modifications of the Cooke type are shown on the right-hand side of Fig. 6, of which the most famous lens is the Zeiss "Tessar," designed by Rudolph in 1902 and improved by Wandersleb in 1907, which is like a Cooke lens with the rear element made into a cemented doublet. This has been made by many firms in many sizes and variations, the maximum aperture being about $f/3.5$. The Ross "Xpres," designed in 1913 by Stuart and Hasselkus may be regarded as a Tessar in which the rear element is a cemented triplet instead of a doublet. Another popular type is the "Aldis" lens, which resembles a Cooke in which the first airspace has been removed by cementing lenses I and II together. The front doublet has no focal power, but is used merely to remove the aberrations of the rear biconvex simple lens.

Attempts have been made to design a lens of the Celor type in which the convex elements are inside, adjacent to the stop, and the concave elements are outside. H. D. Taylor patented such a design in 1912, in which he found it necessary to split each convex lens into two, thus giving twelve glass-air surfaces! A more successful design was the "Plasmat" designed by Rudolph in 1920 and manufactured by the firm of Meyer in Goerlitz. Each element of this lens is like the opposite element of the Planar. This lens has been constructed in apertures up to $f/1.5$ for motion picture purposes.

Telephoto Lenses

A Telephoto lens is simply one consisting of a convex front element and a concave rear element, separated by a considerable distance, so that the second principal point of the whole system is out in front of the convex element. Under these circumstances, the true focal length, which determines the size of the picture, is much greater than the back focus, which determines the "bellows-extension" of the camera. "Telephoto Magnification" is taken as being the ratio of the true focal length to the back focus.

The earliest recorded application[5] of such a lens to photography is the use of a Galilean telescope thrown out of focus, by I. Porro in 1851. This was followed in 1891 by specially designed achromatic negative lenses which were to be used behind an ordinary photographic lens as a "telenegative" lens. Unfortunately, there is a theorem that it is impossible to get a flat tangential field by combining two aplanatic lenses either of which has a curved field. Hence a thin telenegative lens used with a fully corrected positive lens must give a curved field if the final image is to be free from spherical aberration and coma. However, at $f/11$ or less, the telenegative lens was fairly useful, especially as the magnification could be varied by altering the separation of the two systems, but even this led to trouble because the spherical aberration of the negative lens could only be corrected for one separation.

The user of telephoto lenses at first thought he wanted a magnification of at least 8 times, but he gradually realized that 2 or 3 times is all he really needed. Consequently, and with great improvement in the image, Zeiss in 1898 produced a complete telephoto lens, in which the power was $2\times$ to $3\times$ and the aperture $f/6$ to $f/10$. This lens was followed by many others of a similar type, notably the fixed-focus Busch "Bis-Telar" telephoto lens designed by Martin in 1906, which had a magnification of 2 and an aperture of $f/7$; and the Zeiss "Magnar" of Rudolph and Wanderslab, which magnified 3 times at $f/10$.

In 1902, Dallmeyer produced the "Adon," which was a small $2\times$ or $3\times$ afocal Galilean

[5] Waterhouse, Proc. Opt. Convention 1905, p. 115.

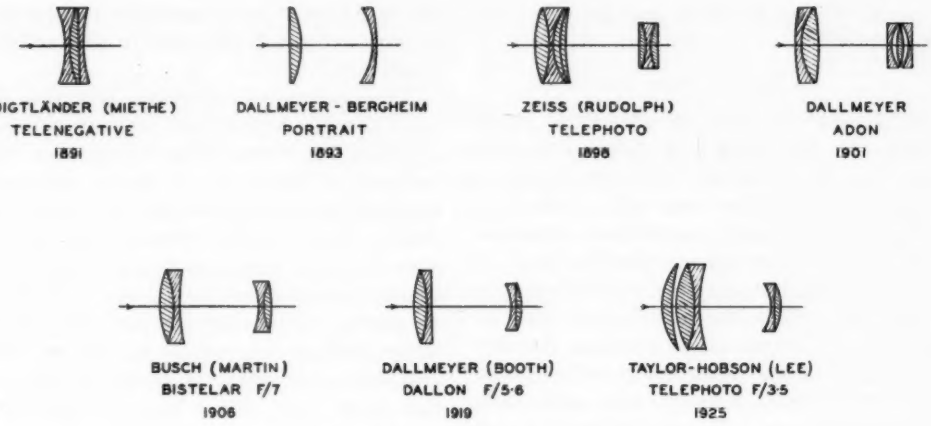


FIG. 7. Some telephoto lenses.

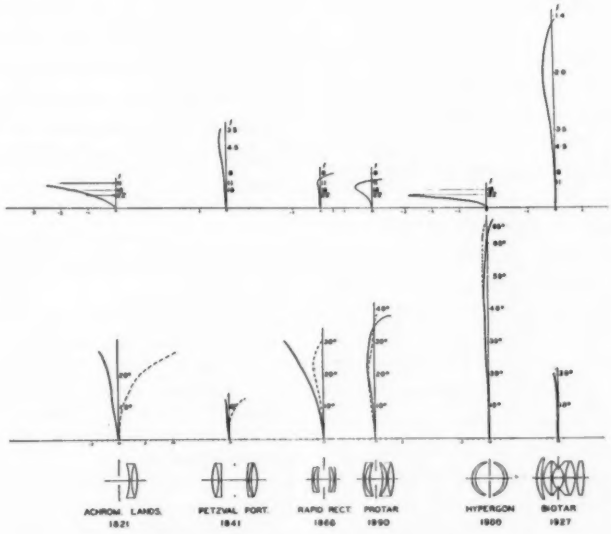


FIG. 8. Aberration and field curves.

telescope made to slip on in front of the ordinary camera lens. Soon it was found that the Adon could with slight alteration be used separately as a telephoto lens, and it was subsequently (1912) enlarged to an aperture of $f/4.5$ or $f/6$ as a separate telephoto system. In 1914, Booth[6] designed the Cooke telephoto lens of aperture $f/5.6$, and $2\times$ magnification, which started the move towards large apertures, and he followed this by the Dallmeyer "Dallon," $f/5.6$, in 1918. Other firms followed this plan, which reached a climax in the Cooke $2\times$ telephoto $f/3.5$ designed by Lee in 1925; and the Cooke Distortionless telephoto $f/5.0$, magnification 2.3, also designed by Lee (Fig. 7).

An interesting use of a reversed telephoto lens placed with the concave element facing the distant object and the convex element towards

[6] Booth, Proc. Opt. Convention 1926, 861–877.

the plate, is seen in Hill's lens for whole-sky photography.[7]

The Future

The possibilities open to a designer in developing new lens types are almost unlimited, except that he must use only the available varieties of glass. Most new types have been introduced in an effort to achieve a greater aperture or a greater angular field or both, to improve the distribution of inevitable aberration residuals, to reduce the number of glass-air surfaces, or to cheapen existing designs. Another potent factor is the avoidance of patents held by other manufacturers. Attempts have been made, also, to improve the uniformity of illumination over the field by shortening the overall length of the lens barrel, or by making the lenses larger in diameter than the greatest stop with which they are used.

An indication of the progress that has been made is seen in Fig. 8. In the upper diagrams, spherical aberration is shown plotted against aperture, and in the lower diagrams are shown the sagittal (~~dotted~~ full) and tangential (full ~~dotted~~) field curves.

The tendency today is towards greater and greater relative apertures, mainly for use in motion picture work where fields of frequently only 8° to 10° from the axis are used. The climax is at present the Zeiss "R-Biotar" with an aperture of $f/0.75$, which covers a 16-mm picture with a focal length of 50-mm, representing an angular field of only 9° from the axis. More rapid photographic materials lead to a demand for faster lenses, because new possibilities in indoor photography under low illumination, etc., become available, which were not contemplated 10 years ago. Moreover, it happens very fortunately that this demand for faster lenses is largely in the motion-picture field, where only very short focal lengths are needed. Now, as the focal length of a lens of given construction is reduced, the aberrations are reduced in proportion, while the permissible *tolerances* remain unchanged. Thus much greater relative apertures become possible in lenses of small size than in similar lenses of larger size.

The zonal residual spherical aberration of these large-aperture lenses often becomes so large that attempts have been made to produce aspheric surfaces to reduce it. A simple and accurate means of generating nonspherical surfaces of any desired form would revolutionize the whole science of lens design, and would make possible simple types of lenses having apertures and fields at present quite unattainable. This would be further helped very greatly by a wider range of available glass types, especially glasses with a high index and also a high V-value, beyond anything at present available in dense barium crown glass. As far as theory is concerned a new calculus is really required, that will enable us easily to take a light-wave of a given form, follow it along to the image, and determine the light distribution in that image in a few minutes! Such a determination, if available, would make interferometer methods of lens testing of much greater value to lens designers than they are at present.

[7] Proc. Opt. Convention 1926, 878.

Bibliography

M. von Rohr. *Theorie und Geschichte des photographischen Objektivs.* (Springer).

W. Merté, R. Richter and M. von Rohr, *Das photographische Objektiv.* (Springer).

J. M. Eder, *Die photographischen Objektive.* (Knapp).

O. Lummer, *Contributions to Photographic Optics,* translated by S. P. Thompson. (Macmillan).

C. Beck and H. Andrews, *Photographic Lenses.* (Beck; and Lund Humphries).

J. Traill Taylor, *The Optics of Photography and Photographic Lenses.* (Whittaker).

Chapters on photographic lenses are also included in the following:

Encyclopaedia Britannica, Eleventh edition, article *Photography* by A. H. Hinton.

A. Gleichen, *Theory of Modern Optical Instruments,* translated by H. H. Emsley and W. Swaine. (H. M. Stationery Office).

C. B. Neblette, *Photography.* (Van Nostrand.)

F. Auerbach, *The Zeiss Works,* translated by R. Kanthack. (W. and G. Foyle).

H. Hovestadt, *Jena Glass,* translated by J. D. and A. Everett. (Macmillan.)

The Effect, on the Latent Image, of Post-Heating and of Melting the Gelatin

A. J. REARDON AND H. P. GRIGGS, S. J., *Saint Louis University, St. Louis, Missouri*

(Received July 17, 1933)

INTRODUCTION

IT is found in general that the application of heat to the photographic emulsion results in an increase in the sensitivity, regardless of whether pre- or post-heating is employed. Toth,[1] Schaumann,[2] Henderson[3] and Acworth[4] agree that previous heating causes a rise, though in some cases it may be small, in the sensitivity of the photographic plate. Kellner,[5] who has made a study of both pre- and post-heating effects on the emulsion, concluded from his experiments that "a rise in sensitivity can come about by heating the plates after exposure to light." This, however, seems to contradict the work of Baekeland,[6] who says that one can partially erase the latent photographic image by keeping the exposed plate for a sufficient time at a suitable temperature (3 days at 50°C). Belliot,[7] whose work is much more recent than that of Baekeland, also concluded that the densities of the previously exposed regions are lowered under heat action, but that at the same time the density of the non-exposed region is increased. Kellner[5] also shows that a selective influence, observed by Osamu Masaki,[8] due to heating during exposure, does not exist, but that this procedure resulted in a decided general increase in sensitivity as was previously established by Abney.[9] Mlle. J. Clavier[10] in her recent work, a study of the influence of temperature on the sensitivity of several kinds of photographic plates exposed to different radiations (4000–5500A), confirms these results.

[1] Toth, Phot. Korresp. S. 205 (1884).
[2] Schaumann, Phot. Wochenblatt, S. 205 (1884).
[3] Henderson, Phot. News, S. 447 (1884).
[4] Acworth, Phot. J., S. 20 (1892).
[5] Kellner, Zeits. f. wiss. Phot. **24**, 63–78 (1926).
[6] Baekeland, J. Phys. Chem. **15**, 314–323 (1911).
[7] Belliot, Comptes Rendus **183**, 1279 (1926).
[8] Osamu Masaki, Chem. Centralbl. (1924); II; (1889). Jap. J. Phys. **2**, 163 (1923).
[9] Abney, Brit. J. Phot., S. 306 (1884).
[10] Mlle. J. Clavier, Comptes Rendus **193**, 235 (1933).

It will be noted that all the experimentation in this field, thus far mentioned, has been concerned with dry heating in ventilated and unventilated ovens. Kellner,[5] who compared plates heated in ordinary air with those heated in air dried in phosphorus-pentoxide, concludes that great dryness brings about a decrease of the post-heating and an increase in the pre-heating exposure effects. In the present paper the authors are concerned with the effects produced by heating in an oven saturated with water vapor.

APPARATUS

The material used in these experiments was limited to Extra Fast, Dry, Undyed Hammer Plates* size 9×12 cm. The source of light was a 21 candle-power 12–16 volt Mazda lamp operated at 12 volts a.c., the current for which was obtained from the city supply by means of a transformer. The lamp was housed in a box and placed 3.9 cm behind an opening 4.2×1.7 cm. This opening was covered with three pieces of glass in tandem (lantern slide thickness) ground on both sides with emery powder. The plate-holder, mounted on a movable frame, was placed 188 cm directly in front of the source of light. Exposures were made, one at a time, by moving the plate across a rectangular opening 0.7×1.2 cm which faced the source of light.

By considering Fig. 1, which shows one of the plates cut in half, it will be noted that there are four strips, each consisting of three exposures. The times of exposure were respectively 6, 30 and 90 seconds. The exact time of exposure was regulated by watching the sweep hand of a Telechron electric clock. The clock and lamp circuits were closed simultaneously by means of a double pole switch.

The source of heat was a 400 watt electric

* We wish to thank the Hammer Plate Co. of St. Louis, Mo. for the plates and for their hearty cooperation.

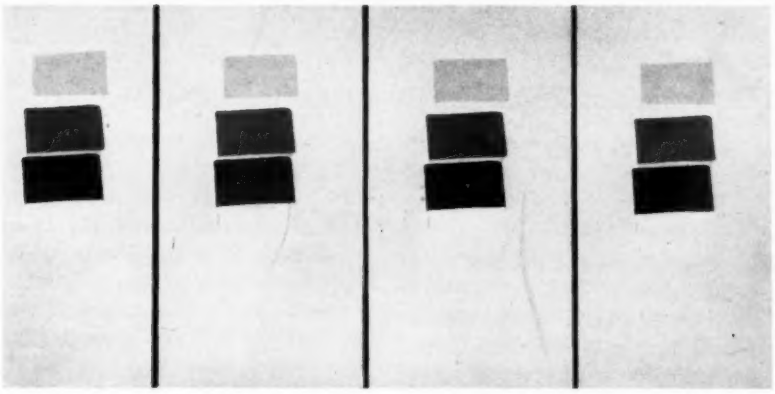


FIG. 1. A reproduction of half of one of the plates. The strips were obtained by means of vertical cuts as shown by the black lines.

water heater connected in series with a variable resistance. A rack which would hold two 9×12 cm glass plates was mounted over the opening of the heater. The lower plate, which served as a support for the exposure strips, rested 3 cm above the top of the heater. The second plate, placed 3 cm above the lower, acted as a protection for the experimental strips against drops of condensed water vapor. A thermometer, fixed at the same height as the rack, and very close to it, indicated the temperatures. A bell-jar was placed over the heater and thermometer, thus completing the oven.

All exposure strips were developed simultaneously for a period of five minutes in the following developer which was kept at a constant temperature of 65°F.

Development Formula	
Water	3632 cc
Hydrochinon	75 g
Sodium sulphite	240 g
Sulphuric acid	3.5 cc
Potassium carbonate	180 g
Potassium bromide	15 g

Acid Hypo Fixing Bath Formula	
Water	240 cc
Sodium hyposulphite	60 g
Water	75 cc
Sodium sulphite	15 g
Acetic acid (28%)	40 cc
Powdered alum	15 g

Uniform development was insured by light brushing.[11] Each strip of exposures was washed for one minute in water and then fixed in the above acid hypo solution. After an interval of 15 minutes the strips were removed and placed in running water where they remained for two hours.

The diffused densities were measured by means of the "Differential Photoelectric Photometer" as described by F. K. Richtmyer.[12]

PROCEDURE AND RESULTS

After a number of preliminary plates had been made we adopted the following procedure. Two plates were used for each experiment. Each plate, as may be seen by referring to Fig. 1, was divided, by cutting, into eight strips of three exposures each (6, 30 and 90 seconds). One strip from each plate was retained as a "standard" without further treatment. The other seven, in one case, were given the following treatment previous to heating in the oven. They were placed, one at a time, in water at a temperature of 45°C for a period of five seconds. Here we must call attention to the extreme care which had to be exercised while handling the strips in the warm water. It was only after many failures that we learned the technique of dipping the strips into the water in such a manner as to retain the emulsion in its original position. Upon removal from the water the strips were allowed to cool and dry thoroughly. The seven strips from the second plate were not given this preliminary water-treatment.

[11] W. Clark, Phot. J. **76**, Feb. (1925).

[12] F. K. Richtmyer, Phys. Rev. **6**, 66 (1915).

The comparative part of the experiment was performed in the following manner. The two plates mentioned above, were subjected to heat by placing two exposure strips (1 treated and 1 non-treated) in the oven simultaneously. The seven pairs of strips were treated, one pair at a time, by heating them for a period of 30 minutes. The oven temperatures were respectively 30°, 40°, 50°, 60°, 70°, 80° and 90°C.

A second experiment, which may be considered the converse of the one just explained, was next performed. Two more plates were exposed and handled in a manner similar to the ones described above up to their exposure to heat. These seven pairs of strips were placed in an oven whose temperature was kept constant at 70°C. The heating-times for successive pairs of strips were 5, 15, 30, 45, 50, 55 and 60 minutes, respectively.

In Fig. 2, which is a graph showing the results obtained by temperature variations, of the warm-water-treated strips, it will be noted that every curve falls below the "standard." Thus it seems that the greater the oven temperature the less is the actual density of the latent image. This is in agreement with the conclusion of Belliot, namely, the greater the temperature, the less would be the heating-time required to obtain a notable decrease in density; although Belliot did not treat his plates with warm water previous to heating.

In Fig. 3, which shows the results of temperature variations with the non-water-treated exposure strips, it will be noted that there are seven curves besides the standard. Curve 7 represents an oven temperature of 90°C. This curve was never obtained with the warm-water-treated group, because at this temperature excessive fogging, as pointed out by Mlle. J. Clavier,[13] together with a flowing of the emulsion, made it impossible to measure the latent image densities. In this figure we note a maximum in the post-heating effect at an oven temperature of 70°C. This is similar to the phenomenon observed by Kellner in his dry oven experiments, from which he concludes that with rising temperature there is an increase in the sensitivity up to 90°C as a maximum; thereafter is noted a decrease in sensitivity, until, with intense heating fogging takes place.

[13] Mlle. J. Clavier, Comptes Rendus **193**, 235 (1933).

Fig. 4 shows that the curves for the warm-water-treated group fall below the standard. Comparing these curves with those of Fig. 2 we again note that at a greater temperature less heating-time would be required to produce the same results, since the densities for all exposures due to the 5 minutes' heating-time at 70°C are lower, compared with the standard, than those due to a half hour of heating at a temperature of 30°C. It may also be noted that the effect of heating at 70°C for 60 minutes is much greater than the effect obtained by heating at 80°C for 30 minutes.

In Fig. 5 we note a maximum in the heating effect similar to the one found in Fig. 3. In this case at some point between 30 and 45 minutes heating-time the increase in density is neutralized; beyond this point we have a decrease in the latent image density. Again, we note that the post-heating effect at 70°C for a heating period of 60 minutes is greater than the effect obtained with a temperature of 90°C and a heating period of 30 minutes.

The three strips whose graphs are exhibited in Fig. 6 were not heated. The curves shown below the standard are the result of warm-water-treatment alone. Curve 1 was obtained from a strip which had been placed in the water for the usual time of 5 seconds, and curve 2 from one that had remained in the water approximately 15 seconds. A longer period of warm-water treatment was not employed since we found, with an increase in time, that it was impossible to keep the emulsion on the glass. It will be noted that these curves have points corresponding to nine exposure times. This number of exposures was used since we desired to see if any change in the phenomenon might occur between the 6 and 90 seconds exposures which might have been missed by using the 30 seconds exposure as the only intermediate.

Discussion

According to Andressen and Lüppo-Cramer,[14] exposed and unexposed silver bromide is reduced at the same rate by the developer. However, when a binding material, such as gelatin, is employed there is a decrease in the rate at which

[14] Andressen and Lüppo-Cramer, Zeits. f. wiss. Phot. **1**, 377 (1903).

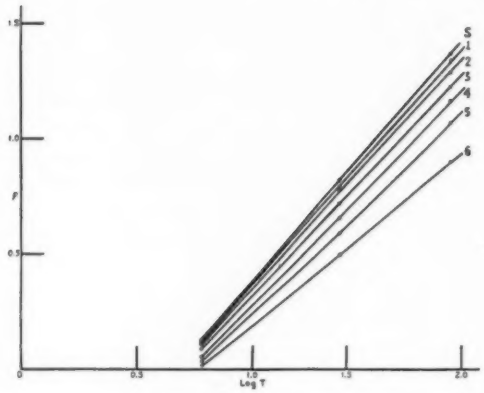


FIG. 2.

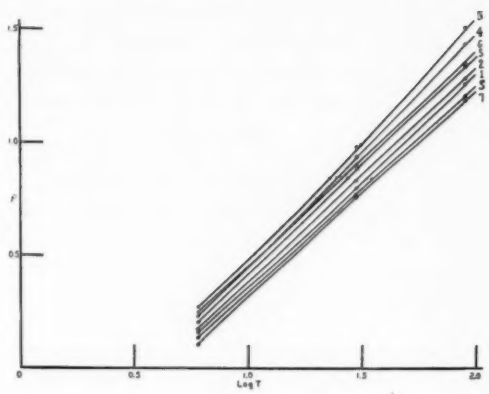


FIG. 3.

FIGS. 2 AND 3. Effect of temperature variations, with constant heating-time, for the treated and non-treated exposure strips. *S* indicates the standard of comparison. Curves 1–7 represent temperatures of 30°, 40°, 50°, 60°, 70°, 80° and 90°C, respectively.

this reduction takes place. In the light of this and the crystal-cracking hypothesis we suggest a possible explanation of the experimental facts noted above. Our experimentation was limited to one type of emulsion. In view of this and of the fact that the effects produced by heating and melting the emulsion, after exposure, are exceedingly complicated, we realize the necessity of further study, especially with regard to different types of emulsions, before definite conclusions may be reached.

The water used in the pre-oven treatment, as described above, was of a temperature sufficient to melt the gelatin. This was confirmed by examinations made outside the dark room. As a result of this treatment, we have a gelatin, mixed

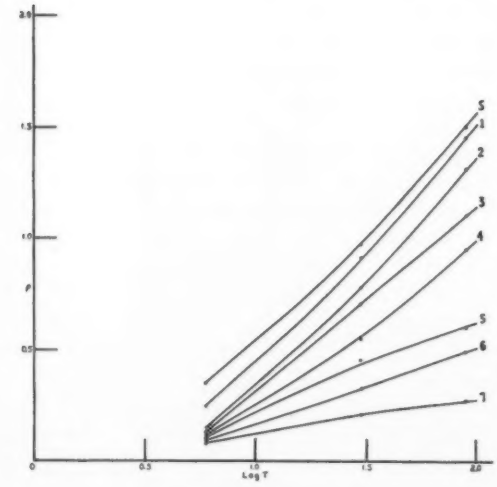


FIG. 4.

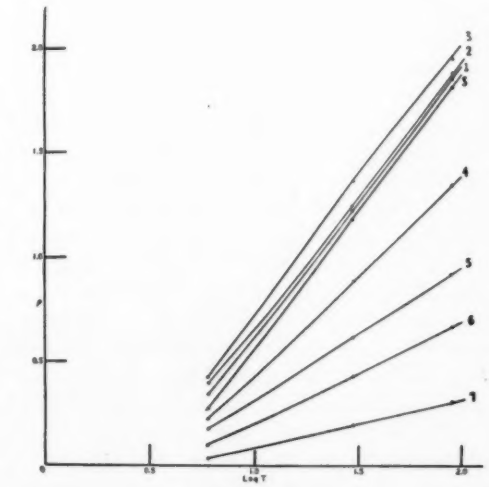


FIG. 5.

FIGS. 4 AND 5. Effect of variations in heating-time, while a constant temperature of 70°C is maintained, for the treated and non-treated exposure strips. *S* indicates the standard of comparison. Curves 1–7 represent heating-times of 5, 15, 30, 45, 50, 55 and 60 minutes, respectively.

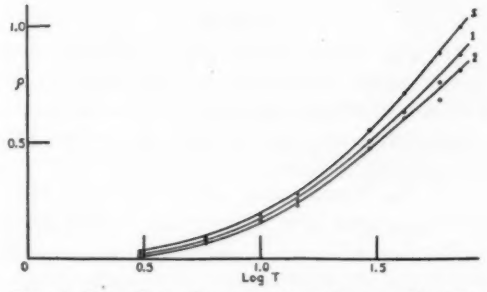
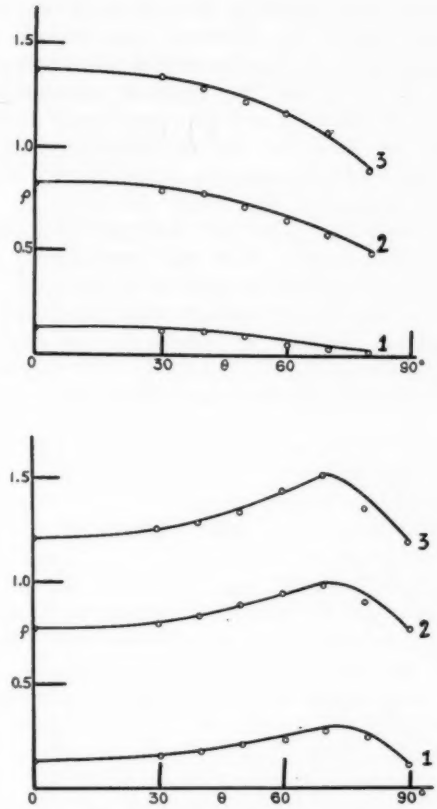


FIG. 6. The effect of warm water alone on the latent image. *S* represents the standard of comparison and curves 1 and 2 warm water treatment periods of 5 and 15 seconds, respectively.

with and expanded by water, which has a greater fluidity than the original gelatin. The cracks developed in the AgBr crystals, by exposure to light, are now covered, to some degree, by this fluid substance. Thus we now have a gelatino-silver-bromide surface instead of a silver-bromide surface. The application of heat forms more cracks in the AgBr crystals but they are rapidly covered by the molten gelatin whose fluidity increases with a rise in temperature or an increase in the heating-time. Referring to Figs. 2 and 4 we are led to believe that the greater the fluidity of the gelatin, the greater the reduction in the density of the latent image.

In the case of the dry plates (Figs. 3 and 5) we have a silver-bromide surface which when it is subjected to heat undergoes strains due to the vibratory motion of the atoms. Up to a certain point additional cracks are formed in the AgBr crystals. In the neighborhood of an oven temperature of 70°C and a heating-time of 30 minutes the gelatin, which is affected by the heat, becomes more fluid, gradually flows over the crystals entering some of the cracks. Thereby the grains are rendered less developable on account of the formation of a gelatino-silver-bromide surface.

Referring to Figs. 2, 3, 4 and 5 we note in particular their fan shaped appearance which seems to indicate that the less the initial density, due to light exposure, the less is the effect due to post-heating. For example, if we plot from Fig. 2, as is shown in Fig. 7, the densities obtained for each of the exposures employed, against their respective temperature variations, we find the slope and rate of decrease for the 90 seconds



FIGS. 7 and 8. Relation between density and temperature variations for the treated and non-treated exposure strips. Curves 1, 2 and 3 represent exposure times of 6, 30 and 90 seconds, respectively.

exposure curve to be greater than for the 6 seconds curve. Similar curves for the dry plates, constructed from Fig. 3, are shown in Fig. 8. The slopes of these curves are about the same up to the melting temperature. Thereafter, the curve with the greater exposure to light dips faster in agreement with the wet plates. By constructing the density *vs.* heating-time curves from Figs. 4 and 5, one finds the same effect in the case of the greater light exposure. It appears, therefore, that the greater and more numerous the cracks in the latent image, due to exposure to light, the more readily they are filled with the fluid gelatin.

The experimental facts mentioned in this paper seem to confirm Kellner's observation of a decrease in the post-heating effects; although our

plates were heated in an oven saturated with water vapor, for different time periods and temperatures. In passing we wish to call attention to the fact that the apparent contradiction between the work of Belliot and that of Kellner may be reconciled in the following manner. Belliot held his plates at a temperature of 140°C for a period of one hour. According to our experimental results and also those of Kellner, we would expect, with that temperature and heating-time, such an effect as he observed.

In conclusion, the authors wish to express their gratitude to Dr. F. E. Poindexter, who directed this experimentation, and to Rev. J. I. Shannon, S. J. for their valuable advice and interest.

Summary

(1) The crystal-shattering hypothesis of the latent image formation, as supported by the pressure effect discovered in our laboratories,[15] is apparently confirmed by the experimental work described in this paper.

(2) The post-heating effect, when using an oven saturated with water vapor is twofold: (a) an increase in density is observed previous to the melting of the gelatin. (b) a decrease is observed once the gelatin has been melted.

---

[15] Poindexter, J. Opt. Soc. Am. **21**, 59 (1930). Maring, S. J. Phys. Rev. **42**, 911A (1932). Also cf. Ny Tsi-Ze and Chien Ling-Chao, Sci. et Ind. Photo., Jan. (1933).

Some Comments on the Ives Theory of Flicker

PERCY W. COBB, *Department of Ophthalmology, Washington University, St. Louis, Missouri*

(Received December 14, 1933)

THE time-relations existing between the stimulation of the eye and the visual response have excited interest to the point that a complete review of the subject would make a considerable chapter. One aspect of this subject has had a twofold interest, that is, the result of recurring or intermittent stimulation. Its original interest was its application, first, it appears, by Talbot[1] as a means of reducing the intensity of a light beam in photometry. Very soon thereafter Plateau[2] brought forward results to show the resultant visual effect of the interrupted beam. This result has been formulated quantitatively in Talbot's law, or the Talbot-Plateau law, which presupposes that the rate of recurrence is sufficiently high so that the visual effect is continuous and free from flicker, and states that the visual result is then the same as that of a continuously acting light, whose intensity is equal to the average of the intensity of the variable or interrupted light. Its most recent and precise verification is no doubt that of Hyde.[3]

From this point of view the frequency of recurrence at which fusion takes place and flicker disappears is of secondary interest, and only to the point of getting rid of the flicker. On the other hand, the critical frequency of stimulation which marks the point of fusion is a datum which has its own interest, by reason of the implications it may have as to the mechanism of the visual process.

Early investigators, as Ferry[4] and Porter,[5] have shown the relation between critical frequency and the intensity of the light, using a cycle of stimulation composed of equal light and dark intervals. Their results have been reduced to formulas, which state in effect that the critical frequency increases as the logarithm of the intensity, or $n=a+b \log I$, in which n is the number of interruptions per second, I the intensity of the interrupted light, and a and b are constants for which Porter gives numerical values.

While a formula of this sort may be convenient as a concise statement of the results actually obtained, it contains no intimation as to the physiological or other processes which bring about those results. It is simply a shorthand substitute for an extended table of data. How misleading a formula may become, especially if it contains the symbol "log," will be evident from a remark made by Troland:[6] "On the assumption that the disappearance of flicker with increasing speed of alternation is due to the reduction of the actual change in sensation below the threshold, the logarithmic law appears to be a corollary of Fechner's law, if we suppose that the change in brightness is proportional to the duration of the individual phase." Now the nearest supposition we can make is that the fluctuation in the effect of stimulation, measured in corresponding terms of stimulus units, is proportional to the intensity, as well as to the duration of the individual phase, whence Weber's law (which I think Troland means here)[7] would be fulfilled without change in speed. The fact that the critical speed rises with increasing intensity is then a distinct refutation of Weber's law for this case.

In recent years H. E. Ives[8] has proposed a

[1] Talbot, *vide* Helmholtz, *Physiological Optics*, English Translation **2**, p. 199.

[2] Plateau, *vide idem*, **2**, pp. 208–209.

[3] E. P. Hyde, *Talbot's Law as Applied to the Rotating Sectored Disk*, Bull. Bur. Stand. **2**, 1–32 (1906).

[4] E. S. Ferry, *Persistence of Vision*, Am. J. Sci. [3] **144**, 192–207 (1892).

[5] T. C. Porter, *Contributions to the Study of Flicker*, Proc. Roy. Soc. **A63**, 347–356 (1898); **70**, 313–329 (1902); **A86**, 495–513 (1911–1912).

[6] L. T. Troland, Trans. Ill. Eng. Soc. **26**, 153 (1931).

[7] The distinction will be found more fully set forth in: P. W. Cobb, *Weber's Law and the Fechnerian Muddle*, Psych. Rev. **39**, 533–551 (1932).

[8] H. E. Ives, (a) *Critical Frequency Relations in Scotopic Vision*, J. Opt. Soc. Am. **6**, 254–268 (1922); (b) *A Theory of Intermittent Vision*, ibid. pp. 343–361. Since these two appear in the same volume, reference has been made largely by page number only.

theory which assumes a physically definite mechanism to account for the variations in the critical frequency of flicker. He proceeds from the widely accepted photochemical theory, and bases his reasoning on the behavior of a photoelectric cell with a liquid electrolyte. Such a model is, however, not necessary, since all essential relations are provided by a hypothetical retinal cone, let us say, according to the scheme made definite by Hecht,[9] with some additions not mentioned by him.

We suppose in a receptor a substance S which, under the action of light, is decomposed into two others, P and A, of which P alone is the actual nerve-stimulating substance, according to the reaction: $S=P+A$; the velocity of the process being proportioned to $a-x$, the concentration of S at any instant, and to the intensity I of the light acting:

$$v_1=k_1I(a-x). \tag{1}$$

At the same time, a resynthesis of S is taking place according to the formula: $P+A=S$. This reaction will have a velocity proportional to the product of the concentrations of P and A, which are each, being equal, represented by x:

$$v_2=k_2x^2. \tag{2}$$

The rate of change in concentration of P at any instant will be:

$$dx/dt=v_1-v_2=k_1I(a-x)-k_2x^2. \tag{3}$$

It will be seen that a is here the sum of the concentrations of S and P, or the theoretical maximum of either one.

If now a light of constant intensity I' acts on such a chemical cell, eventually a stationary state will be reached:

$$k_1I'(a-x)-k_2x^2=0 \tag{4}$$

at which point the concentration of P will be x', the positive root of the quadratic equation. If now the stimulus is suddenly changed from I' to I, we have:

$$dx/dt=k_1I(a-x)-k_2x^2, \tag{5}$$

in which, at the very first instant, x is still equal to x'.

By subtracting, (4) from (5) yields:

$$dx/dt=k_1(I-I')(a-x'). \tag{6}$$

That is, the rate of change in x at any instant is proportional to two things: the acting stimulus, less the stimulus with respect to which the system is quantitatively in the stationary state; and to the sensitivity of the system $(a-x')$ at that instant. If the time is so short that the state of the system, measured in terms of corresponding or equilibrium stimulus, changes very little as compared with the difference $I-I'$, then also will the change in x be small and the following stands as a close approximation for sufficiently brief finite times:

$$dx/dt=c(I-I'). \tag{7}$$

With these provisions, the effect of an interrupted stimulus would be a saw-tooth variation in x, as it rises and falls uniformly during the light and dark periods respectively, which is exactly the relation Ives describes (pp. 349–350) for the photoelectric cell. It should be further pointed out that there are many more or less simple physical systems of which such a relation will be true. It is only necessary that the range of variation be sufficiently small with reference to the range given by the equilibrium states corresponding with the extreme values of the exciting factor. It will be shown later that certain of Ives' data do not satisfy this requirement. Also, and incidentally, it may be shown with little trouble that Eq. (7) is deducible from the Talbot-Plateau law, and is therefore confirmed by the experimental fact that the law holds very precisely.

Further, Ives proposes that the substance P must be carried by diffusion from its site of production to the site of stimulation, according to a law of diffusion stated:

$$dx/dt=K(d^2x/dh^2)-\mu x, \tag{8}$$

in which t is the time, h the distance of diffusion, μ the coefficient of loss of the material P, and K the constant of diffusion.

From this point the discussion is helped by the fact that it is mathematically possible to express any recurrent pattern of stimulus-intensity by means of a Fourier series.

[9] Selig, Hecht, see for example, J. Opt. Soc. Am. **18**, 276 (1929); or Ergebnisse der Physiologie, **32**, 256–257 (1931).

The process may then be followed through mathematically. In what follows we will designate the separate stages, for convenience: Stage I, the equation of the stimulus itself as a function of time; Stage II, the statement of the initial photochemical effect; and Stage III, the result after diffusion has taken place to a distance h, where, say, the impulse first becomes a characteristically nervous one. The transition from I to II is thus the first step, as designated by Ives (p. 346), and from II to III the second step.

For an abruptly interrupted and abruptly resumed stimulus, with equal intervals on and off, the value of the stimulus at any time t is:[10]

$$I_t = I/2 + (4Ia/\pi)(\cos \omega t - \tfrac{1}{3} \cos 3\omega t + \tfrac{1}{5} \cos 5\omega t \cdots). \quad (9)$$

Since the rate of change of x is proportional to $I_t - I'$ and I' is taken as constant and equal to average I, the variable portion of the expression for x at Stage II will be the variable portion of I_t integrated with respect to time and multiplied by a constant c:

$$x_t = x_{av} + (4Ia/\pi)(c/\omega)(\sin \omega t - (1/9) \sin 3\omega t + (1/25) \sin 5\omega t \cdots). \quad (10)$$

The expression for the value of x at any time, after diffusion to a distance h, obtained by applying the general solution of (7) to the foregoing is:

$$x_h = x_{av} \exp[-(\mu/K)^{\frac{1}{2}} h] + (4Iac/\pi\omega)[\exp\{-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\} \sin\{\omega t - (h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}-\mu]^{\frac{1}{2}}\} - 1/9 \exp\{-(h/(2K)^{\frac{1}{2}})[(\mu^2+9\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\} \sin\{3\omega t - (h/(2K)^{\frac{1}{2}})[(\mu^2+9\omega^2)^{\frac{1}{2}}-\mu]^{\frac{1}{2}}\} + \cdots]. \quad (11)$$

Following this statement Ives says (p. 352): "Now if the amplitude of this function is small, the part contributed by the periodic terms after the first, involving higher multiples of ω in the exponential term, may be neglected."

In the earlier paper however Ives gives results which make this assumption utterly untenable for a certain range of the data presented. Referring to p. 263, Fig. 4 of that paper, the upper two curves in the left-hand section, reproduced here in Fig. 1: of these, the upper curve represents the conditions just discussed, and the curve next below represents the results obtained by using, in place of abrupt transition, a stimulus of simple sinusoidal pattern whose equation is:

$$I_t = I/2 + Ia \cos \omega t, \quad (12)$$

which, by the same reasoning as before, yields at Stage II:

$$x_t = x_{av} + (Iac/\omega) \sin \omega t \quad (13)$$

and at Stage III:

$$x_h = x_{av} e^{-f} + (Iac/\omega) e^{-f'} \sin(\omega t - f''), \quad (14)$$

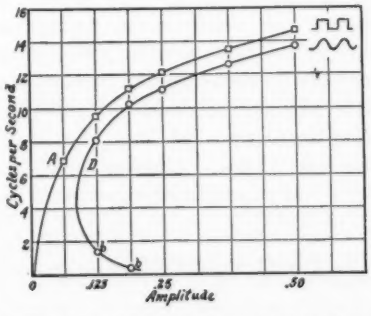


FIG. 1. The data of Ives. The speed, plotted against the amplitude of the stimulus fluctuation: *A*, for a stimulus abruptly interrupted and resumed, with light and dark phases of equal duration; *D*, stimulus intensity a simple sine function of the time. The maximum amplitude, 0.5, occurs when the stimulus varies from full intensity to zero.

[10] It should be explained at this point that while this discussion aims to give a fair exposition of Ives' theory, it departs somewhat from his exact form. Also, unfortunately, certain material errors have crept into his text, to which attention is here called. First, ω, from its position in all his expressions cannot be the frequency in cycles per second as he says (pp. 265 and 348), but must be 2π times that frequency. Further, the parenthetic portions of Eqs. (2), p. 265, and (21), p. 351, should be identical and are not; the former is correct. In Eq. (6), p. 265, the signs of alternate terms within the parenthesis, beginning with the second, should be negative. In Eq. (24), p. 352, first exponent should read $-x(\mu/K)^{\frac{1}{2}}$; in both succeeding terms the first of the pair of braces should follow "sin" and precede ωt and $2\omega t$, respectively, instead of preceding "sin"; and similarly with respect to the sine term in Eq. (25), p. 352.

in which

$$f=(\mu/K)^{\frac{1}{2}}h, \tag{14a}$$

$$f'=(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}, \tag{14b}$$

and

$$f''=(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}-\mu]^{\frac{1}{2}}. \tag{14c}$$

Thus if we neglect all periodic terms but the first, the two sets of results should be identical, except that in the first case the amplitude at Stage I is greater by the factor $4/\pi=1.273$. This factor follows through to Stage III. Accordingly, the two curves are reproduced here again in Fig. 2, with the abscissae (which are plotted as

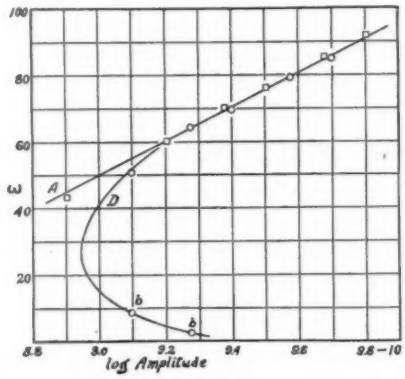


FIG. 2. The same data as Fig. 1, plotted against the common logarithm of the amplitude of the first periodic term of the Fourier series which represents the stimulus variation. The speed ω is 2π times the number of cycles per second.

logarithms) for the upper (square-topped stimulus) curve increased by that factor, so the "amplitude" for either curve represents the amplitude of the first or only periodic term of the expression for the stimulus, or relative amplitude of the effect throughout to Stage III.

Ives explains (pp. 262–263) that with the sinusoidal variation of stimulus it is possible, especially at lower amplitudes, to rotate the disk so slowly that no change whatever is perceived, simply because the rate of change of stimulus is below the threshold. Thus at amplitude 0.125 and 0.1875 there are two points plotted, the lower one *b* at which flicker begins with increasing speed, the upper one where it ceases.

In their upper portions the curves bear out the theory well, but only insofar as they fall together when plotted against the relative amplitudes of the first periodic term of the equation for the stimulus. One must, in any case, ask himself what is happening at the lowest point *A* on the curve for the square-topped stimulus, since at the corresponding point for the sinusoidal stimulus the fluctuation is too small to yield flicker under any conditions, or at the very best at a much lower rate. There is no answer to this except that according to the diffusion theory, which we are discussing, the first periodic term must be failing of effect, as it does in the *D* series, and that the second and possibly others must be involved in the case of the square-topped stimulus.

The reply to this might be made that, after all, the difference at the point *A* is small and practically negligible. This consideration would have more weight did not the two curves unfortunately, fall together so nicely in the region above that point. Also, common sense explains the divergence. No matter how far the speed is reduced, it is obvious that with the abrupt change, there will of necessity be a marked sense of change for the observing eye. Or, in mathematical language, there will always be, in the Fourier series, a periodic term of sufficiently high frequency to produce flicker. The importance of this will be more evident if we turn to other stimulus patterns.

If instead of using equal light and dark intervals in the case of the square-topped stimulus, we depart from that ratio, the Fourier series for the stimulus (Stage I) will be:

$$I_t=I_{av}+(4Ia/\pi)(\sin\pi\varphi\cos\omega t+\tfrac{1}{2}\sin 2\pi\varphi\cos 2\omega t+\tfrac{1}{3}\sin 3\pi\varphi\cos 3\omega t+\cdots), \tag{15}$$

in which φ is the ratio of the light phase to the whole cycle. It will be noticed that the term containing $2\omega t$ appears here, as it does not in the case of equal light and dark intervals where $\varphi=\frac{1}{2}$. In the latter case the terms containing even multiples of ωt drop out, since their coefficients, $\sin 2\pi\varphi$, $\sin 4\pi\varphi$, etc., all become zero. However, as φ approaches its maximum and

minimum values, unity and zero, the second periodic term containing $2\omega t$ becomes significant, and then of increasing importance, first, because in the final equation, corresponding to Eq. (11) it is 2ω and not 3ω which is involved in the negative exponent, and further, because (especially when φ approaches closely to unity or zero) the other factor in the coefficient of the second periodic term containing $2\omega t$ will be $\frac{1}{4}$ sin $2\omega t$, or nearly one-half that of the first term (sin ωt), instead of nearly one-third as in the case of equal light and dark intervals.

A point which Ives' treatment fails to explain is the return of the curve for the sinusoidal (D) series upon itself, at b, b, Fig. 2. From Eq. (14) it appears that the amplitude of the fluctuating term at Stage III is the coefficient of the sine:

$$(Iac/\omega) \exp \{-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\} \quad (16)$$

and the rate of change in x, which Ives finally (pp. 352–353) takes to be the determinant of the flicker-point, is the derivative of the sine term with respect to t, in which the sine will become cosine, and ω will disappear from the denominator. The maximum rate of change in the cycle is then the coefficient of the cosine or:

$$Iac \exp \{-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\}. \quad (17)$$

Now since the exponent can only increase in the negative sense with ω, and the above expressions can therefore do nothing but decrease with increase in ω, the "returning" values, b, b, Fig. 1, are completely unexplained by Ives' treatment. The difficulty lies in the transition from Stage I to Stage II, as shown in Eqs. (9) and (10), where it is assumed by Ives that the rate of change in x is proportional to the deviation of I_t from its own average value. This may *only* be assumed to be approximately true as long as the value I', that is, the value of stimulus with which x at that instant would be in equilibrium, deviates from its average by an amount very small *as compared with the corresponding deviations of the actual stimulus I_t, regardless of how small these deviations may be in themselves.*

The discrepancy may be avoided by the following procedure. While we cannot wholly accept the photochemical cell as described above (Eqs. (1) to (5)) as yielding the true values of x with respect to I[11] for the equilibrium state, we may with considerable confidence assume this relation to be linear for a limited region, especially when we are dealing with oscillations which arise from low amplitudes of fluctuation of the original stimulus at Stage I:

$$x = AI' + B, \quad (18)$$

A and B being constants. From this we may derive a relation between x and I' which, substituted in (7) as indicated, will yield:

$$(dx/dt) = A\,dI'/dt = c(I_t - I'), \quad (19)$$

which, in turn, by first disregarding x, integrating, then replacing I' by the value of x given in (18), and finally introducing the unknown average value of x as the constant of integration, yields:

$$x_t = x_{av} + AIa(k/(\omega^2+k^2)^{\frac{1}{2}}) \sin(\omega t - \tan^{-1}\omega/k) \quad (20)$$

in which $k = c/A$.

Now if this is followed through to Stage III, then differentiated, the maximum rate of change will be found to be:

$$AIak(\omega/(\omega^2+k^2)^{\frac{1}{2}}) \exp\{-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\}, \quad (21)$$

which differs from (17) in one essential particular only, that it contains the factor $\omega/(\omega^2+k^2)^{\frac{1}{2}}$. The effect of this is, that at low speeds, when ω is small as compared with k, its effect in the radical of the denominator becomes comparatively insignificant, and ω in the numerator is left as an important factor in the expression, instead of k,

[11] As a matter of fact, the quadratic relation between x and I' in Eq. (4) rather closely approaches a straight-line relation between x and log I' for a limited region between $x/a = 0.35$ and $x/a = 0.75$, thus approximately fulfilling Weber's law on the assumption that a constant difference in x is the condition of the difference threshold. The question whether the final criterion of flicker should involve the average value of x in the denominator would seem, by this, to be tentatively disposed of in the negative.

or c as in (17). On the other hand, at high speeds, as ω rises to overshadow k in the radical, the fraction approaches unity in value and the expression (21) comes to have the same form as (17), which, except for the exponential factor, is independent of ω. Thus (17) is a special case of (21), which is the more general expression.

Thus the peculiar course of the curve for the D (sine-wave stimulus) series finds a mathematical interpretation, since at low speeds, the oscillation at Stage II, and also at Stage III, increases with the speed, and the lower limit of flicker in the D series is thereby explained. This feature does not appear in the curve for the A (square-topped stimulus) series, simply because, as will readily be seen from (9), there are terms in which the frequency has values which are 3 times, 5 times, etc., its value in the first with respect to which the fraction $\omega/(\omega^2+k^2)^{\frac{1}{2}}$ approaches unity, and for which the expression (21) therefore approaches expression (17) in value.

From the data, D series, Fig. 1, the value of k would appear to be about 3 or 4, when ω is 2π times the number of cycles per second. That is, this value is found to be fairly satisfactory for the points b, b, but a closer approximation cannot be reached until it is more definitely ascertained what the factor is that degrades the oscillation at speeds higher than this, say from about 10 cycles per second up, when ω is equal to about 60 or more. In this range, with k as high as 5, the value of the fraction $\omega/(\omega^2+k^2)^{\frac{1}{2}}$ becomes $60/(3625)^{\frac{1}{2}}$ or 0.9965. In this region then the effect of the correction introduced by the modified procedure of Eqs. (19), (20) and (21) may be neglected, as less than 0.4 percent, and the upper course of the curve may be discussed from the standpoint of (17). It is shown in Fig. 2 that the results of the A and of the D series when plotted against the relative amplitudes of the first or only terms of their respective Fourier series at Stage I, agree very closely in this upper region, and they display as well a very closely linear relation to the logarithm of the amplitude, for which the following equation has been derived from the plot:

$$\omega = 101.3 + 22.3 \ln a. \qquad (22)$$

If now we consider the hypothesis of Ives, that a certain minimal value of the "slope" or first derivative of the oscillation at Stage III is the condition of flicker, we have in the above, taken in connection with (17), a means of testing the theory as far as diffusion is concerned, in the upper reaches of the curves, where it may be assumed with certainty that one, and only one, periodic term is involved.

Examination of

$$Iac \exp\{-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}\} \qquad (17)$$

shows the rate of change at Stage III, and consequently the tendency to flicker, increases, first, directly as the amplitude a of the stimulus at Stage I. Since the critical speed does not keep up with this, the theory leaves only the diminishing value of the exponential factor in (17) to explain the deviation. This will be more obvious if we put (17) equal to a small minimal value r, throw the equation so formed into logarithmic form, and then for ln a substitute its empirical value $(\omega-101.3)/22.3$ found by solving (22) and collect the terms in ω:

$$\ln(r/Ic)+4.54 = 0.0448\,\omega-(h/(2K)^{\frac{1}{2}})[(\mu^2+\omega^2)^{\frac{1}{2}}+\mu]^{\frac{1}{2}}, \qquad (23)$$

in which the second member must be a constant since the first contains no variables.

The last term in this equation is the exponent, which is the natural logarithm of the factor of degradation of the fluctuation in the second step from Stage II to Stage III, but which says nothing whatever about the similar degradation during the first step from Stage I to Stage II, which is represented by the presence of the constant c.

The values of the exponents in Table I have

TABLE I. *The values of the exponents in the factor representing the decrease during the theoretical process of diffusion for stated values of the speed ω, and for selected values of the coefficient of loss μ; and the degree of degradation corresponding to each exponent, which is simply e with the exponent applied in the positive sense.*

	Exponent			Degradation		
μ	$\omega=0$	$\omega=60$	$\omega=90$	$\omega=0$	$\omega=60$	$\omega=90$
0	0	5.98	7.32	1.0	395	1520
1	0.78	6.08	7.43	2.2	437	1680
2	1.11	6.17	7.52	3.0	478	1850
5	1.79	6.47	7.81	6.0	649	2470
10	2.64	7.03	8.38	14	1130	4370
20	4.10	8.36	9.71	60	4280	16400
50	9.07	14.52	15.86	8600	2020000	7730000

been worked out from the above as follows: The two values of ω, 60 and 90, were selected as fairly well representing the empirical straight-line relation between ln a and ω as shown in Fig. 2, and various representative values of μ were selected. In each case the values 60 and 90 were substituted for ω in the final term of (23), the first factor of that term $h/(2K)^{\frac{1}{2}}$, being adjusted to make the difference between the two values of the exponent equal to the difference between the corresponding two values of 0.0448ω, that is, equal to 30×0.0448 or 1.345. The table therefore gives the values which the exponent must have at $\omega=60$ and 90 in order that the theoretical equation shall satisfy the empirical findings at those two points. These two values should then fairly well represent the part of the curve in Fig. 2 which closely follows the logarithmic relation given in (22). Under $\omega=0$, the exponent similarly represents the degradation of the constant effect at Stage II, and is the exponent in the coefficient of the first or non-periodic term in the equation for Stage III.

In the table "degradation" simply means the ratio of "input to output" during the second step of the process, between Stages II and III, and is numerically equal to e raised to the power indicated. For example, if it is equal to 1, there is no decrease; if it is equal to 395 it means that the magnitude of the stimulus effect is reduced at Stage III to 1/395 of its value at Stage II; and so on.

If, now, we wish to evaluate μ so that the exponent is a linear function of ω, as Ives (p. 354) suggests, we will have to make ω/μ equal to about 2, which would give μ a value of 30 to 45. A glance at the lower portion of the table will indicate that the loss is overwhelmingly great for these values, not only with respect to the amplitude of oscillation of the periodic component for the range of speed considered, but equally well for a constant stimulus of the average intensity. It is hardly conceivable, for instance, that a retina which loses more than 98 percent of the effective product of its photochemical activity ($\mu=20$) would be the one which has survived in the age-long process of evolution.

We are not helped much if we make μ much smaller, for then we leave the straight-line relation between ω and the exponent, and get a result so near to the square root of ω that we might as well, as far as fitting the empirical data is concerned, drop μ altogether, that is, put it equal to zero which will make the exponent proportional to the square root of ω. Ives considers this in his earlier paper (p. 267), and comes to the same conclusion as the one above (same page, footnote 9), which is that the amplitude at the third stage, in order to fit the data, has to be placed incredibly low. Furthermore, to substitute ω for its square root, as he does, is merely to discard the theory. This, however, relates to the earlier paper only. The later one finally leaves it in much worse condition, as Table I shows.

On the whole, it does not seem that the physical process of diffusion in the second step of the visual process can be used to account for the relation between frequency and the amplitude of the stimulus pattern. It is, of course, a matter of opinion as to whether the end-effect is or is not too small to be the threshold of effectiveness, but with the least perceptible difference between two constant adjacent stimuli established at about one percent, it scarcely seems possible that 1/15 of this difference, to say nothing of the degradation in the first step implied in the constant c, can be the threshold when applied at very brief intervals to an identical area of the retina.

In any case, for the "peculiar" behavior of the sinusoidally fluctuating stimulus in showing a flicker-range of speed, below which and above which flicker disappears, the amendation of the theory here given (Stage I to Stage II, Eqs. (19) and (20)) offers a mathematical explanation. In pursuing the diffusion theory further, it will no doubt be best to look upon the sine-wave stimulus as the logical place to begin since, unlike the others, its initial form involves one fluctuating component only, which retains its period and fundamental character throughout, while its coefficient changes as it passes through the two steps of transformation. In the case of the abruptly interrupted or square-topped stimulus, we have seen that beyond doubt periodic terms other than the primary one are effective under some conditions.

The foregoing discussion deals wholly with variations in the amplitude of the fluctuating stimulus, and says nothing about intensity. It has been said above that the intensity law has

been worked out empirically by Ferry, and by Porter, and that a mathematically simple relation has been found: fusion-frequency increases in a linear relation with the logarithm of the stimulus-intensity. It has also been said above that this is merely an empirical relation and, as it stands, it has not enabled any one to come any nearer to a physiological analysis of the process than before. It is more than likely that the constants of the assumed photochemical cell, described herein in Eqs. (1) to (5), change when the level of stimulation is shifted. The description given includes no account of possible losses of the substances S, P and A, nor of their possible elaboration *de novo* from simpler chemical materials, both of which would alter the constants of the photochemical cell. Such being the case, the intensity law has to be left alone for the present, and its operation avoided in experimental work which aims to elucidate other phases of the flicker question. Ives has accomplished this elimination by working with blue spectral light of low intensity, under conditions within which it has been shown that there is no change in fusion-frequency with change in the intensity.

It would be appropriate here to suggest another method of eliminating the intensity factor. That is, whatever form of disk or stimulus-cycle is used, to change the intensity of the stimulus to such a degree as is necessary in order that its average value according to Talbot's law shall always be the same at speeds above the critical frequency. In other words, one adopts a standard level of average stimulus for the eye, and then, for the cycle being used at the time, so adjusts the intensity of the original stimulus prior to the disk, that the eye is subject to the standard intensity. In the first place this technique presents no great difficulties. In the second, it is safe to assume, under these conditions, that just above and just below the critical frequency, where one is working to obtain the data, the effect of the stimulation will always be the same as that of the standard, or very close to it; and that hence the assumed photochemical cell will have no occasion to change its constants in order to adapt itself to different intensities of stimulus.

At the worst, there appear no objections to this procedure that do not equally well apply to that of Ives. In its favor is to be mentioned the fact that it is applicable to any level of stimulation, and is not limited to scotopic vision as the other is.

Summary

(1) It is pointed out that a merely empirical formula may be useful as a concise statement of the data, but that it sheds no light on the physiological facts of the case.

(2) The Ives theory of flicker is described and followed through mathematically for illustrative cases.

(3) His assumption that the first periodic term in the Fourier series which describes the stimulus cycle is the determinant of the critical frequency is disproven by results obtained by him with a simple sinusoidally fluctuating stimulus, as compared with those from an abruptly changing or square-topped stimulus of equal light and dark intervals. The case appears to be more serious with certain other forms of stimulus cycle.

(4) This suggests that the analytical study of flicker should logically begin with the simple sinusoidal stimulus, since this contains one and only one periodic term, and the possible complication of the second and following terms is avoided.

(5) A modification on the quantitative treatment of the first step of the process (the transformation from light stimulus to initial photochemical effect) is proposed, which explains the fact that the sinusoidal stimulus yields a lowe r as well as an upper critical speed.

(6) The hypothesis of diffusion, as an explanation of the relation of amplitude to critical frequency shown by the data, is quantitatively examined. It is concluded that the variations in the critical speed are too great to be explained by the diffusion theory without postulating losses which leave the surviving remnant of the fluctuation, or even of the constant term, in the effect too small to be credible.

(7) The known relation between stimulus intensity and critical frequency remains an empirical one.

(8) A method of avoiding the operation of the intensity law, in investigating other phases of flicker, is suggested. This method permits of work under high intensities as well as under the conditions of scotopic vision.